Challenges and prospects of probing galaxy clustering with three-point statistics

Alexander Eggemeier

Department of Physics and Astronomy
University of Sussex

This dissertation is submitted for the degree of
Doctor of Philosophy

June 2018

For my parents.

Declaration

I hereby declare that this thesis has not been and will not be submitted in whole or in part to another University for the award of any other degree. Chap. 4 to 7 have been undertaken in collaboration with other researchers, and the first three of these chapters have been published as

- Eggemeier, Battefeld, Smith & Niemeyer. *MNRAS*, 453:797
- Eggemeier & Smith. *MNRAS*, 466:2496
- Byun, Eggemeier, Regan, Seery & Smith. *MNRAS*, 471:1581

I identify my personal contributions to each of these projects at the end of the relevant chapters.

Alexander Eggemeier
June 2018

Acknowledgements

Needless to say, this thesis would not exist without the support of many different people. First, and foremost, I would like to express my biggest thanks to my supervisor, Robert Smith. You have kept me motivated throughout my entire Ph.D., and through your always careful guidance, assistance and our countless conversations I was able to develop into the scientist I am today. Thank you so much, Robert!

I owe a similar amount of gratitude to Román Scoccimarro, who acted as my second supervisor during my six-month stay at New York University. Thank you for giving me the opportunity to work with you and thanks for all the valuable skills I undoubtedly learned from you. It was great fun!

I consider myself extremely lucky to have been surrounded by so many amazing friends. Coming to Brighton about four years ago without knowing anybody seemed like a daunting endeavor, but meeting you guys turned it into a truly wonderful experience. Particular mention deserve my friends since pretty much the first day and long-term housemates Benoît, Dániel and Ridwan. Your friendships and all of our shared memories are extremely precious to me and I think there is no way I could do them justice here. The same goes for Heena, joining us for our first two years in THOA, and our two more recent housemates Jacob and Panka. Thanks to you all for your unconditional support and for keeping me from working too much.

Another big thank you to the not yet mentioned people of our Ph.D. year: my office mate Hannah, as well as Lucía, Steve and Zé — I am so glad I was able to share the entire journey with all of you. I would also like to thank the rest of my office mates throughout the last four years for making the daily work with its ups and downs so much more enjoyable: Antonio, Azizah, Boyd, David, and more recently, Dan, Lucas, and Robin. Finally, I cannot stress enough how great the entire Sussex Astronomy Department has been and people I have not mentioned so far include: Antony, Ashwin, Carlos, Chris, Ciaran, Elisa, Jesús, Julien (2x), Mateja, Michele, Michaela, Pippa, Rose, Sam, Scott, Sean, Sunayana, and in particular my collaborators David, Donough and Joyce. Thanks for all the time we spent together, the interesting discussions we had, the Squash games, and for putting up with my slow eating habits during lunch breaks. I would also like to thank my NYC crew, consisting of Andrej, Chang, Doycho, MJ and Ruben, for making my stay in the department and in the city so great.

I am deeply grateful to Lucas, Robert and Ridwan for reading parts of my thesis and providing me with very helpful comments to improve its content and appearance. In particular, I would like to thank Ridwan for offering to go through the effort of changing my referencing style.

Last, but far from least, thank you to my family — especially, my parents, as well as my brothers Johannes and Benedikt — you are the real reason why I got to where I stand now. Without your support and love, not just during the last four years, this would have been impossible. I dedicate this thesis to you.

Abstract

In this work we explore three-point statistics applied to the large-scale structure in our Universe. Three-point statistics, such as the bispectrum, encode information not accessible via the standard analysis method — the power spectrum — and thus provide the potential for greatly improving current constraints on cosmological parameters. They also present us with additional challenges, and we focus on two of these arising from a measurement as well as modelling point of view.

The first challenge we address is the covariance matrix of the bispectrum, as its precise estimate is required when performing likelihood analyses. Covariance matrices are usually estimated from a set of independent simulations, whose minimum number scales with the dimension of the covariance matrix. Because there are many more possibilities of finding triplets of galaxies than pairs, compared to the power spectrum this approach becomes rather prohibitive. With this motivation in mind, we explore a novel alternative to the bispectrum: the line correlation function (LCF). It specifically targets information in the phases of density modes that are invisible to the power spectrum, making it a potentially more efficient probe than the bispectrum, which measures a combination of amplitudes and phases. We derive the covariance properties and the impact of shot noise for the LCF and compare these theoretical predictions with measurements from N-body simulations. Based on a Fisher analysis we assess the LCF's sensitivity on cosmological parameters, finding that it is particularly suited for constraining galaxy bias parameters and the amplitude of fluctuations. As a next step we contrast the Fisher information of the LCF with the full bispectrum and two other recently proposed alternatives. We show that the LCF is unlikely to achieve a lossless compression of the bispectrum information, whereas a modal decomposition of the bispectrum can reduce the size of the covariance matrix by at least an order of magnitude.

The second challenge we consider in this work concerns the relation between the dark matter field and luminous tracers, such as galaxies. Accurate knowledge of this galaxy bias relation is required in order to reliably interpret the data gathered by galaxy surveys. On the largest scales the dark matter and galaxy densities are linearly related, but a variety of additional terms need to be taken into account when studying clustering on smaller scales. These have been fully included in recent power spectrum analyses,

whereas the bispectrum model relied on simple prescriptions that were likely extended beyond their realm of validity. In addition, treating power spectrum and bispectrum on different footings means that the two models become inconsistent on small scales. We introduce a new formalism that allows us to elegantly compute the lacking bispectrum contributions from galaxy bias, without running into the renormalization problem. Furthermore, we fit our new model to simulated data by implementing these contributions into a likelihood code. We show that they are crucial in order to obtain results consistent with those from the power spectrum, and that the bispectrum retains its capability of significantly reducing uncertainties in measured parameters when combined with the power spectrum.

Table of contents

List of figures

List of tables

Nomenclature

Acronyms

ΛCDM	Λ Cold Dark Matter
4MOST	4-Metre Multi-Object Spectroscopic Telescope
ALCF	Anisotropic Line Correlation Function
AP	Alcock-Paczyński
BAO	Baryon Acoustic Oscillations
BOSS	Baryon Oscillation Spectroscopic Survey
CIC	Cloud-in-Cell
CMB	Cosmic Microwave Background
COBE	Cosmic Background Explorer
DES	Dark Energy Survey
DESI	Dark Energy Spectroscopic Instrument
EFT	Effective Field Theory
FFT	Fast Fourier Transform
FoG	Fingers-of-God
GAMA	Galaxy and Mass Assembly
GLCF	Galaxy Line Correlation Function
GRF	Gaussian Random Field
HOD	Halo Occupation Distribution
ILCF	Isotropic Line Correlation Function
LCF	Line Correlation Function
LPT	Lagrangian Perturbation Theory
LSST	Large Synoptic Survey Instrument
PDF	Probability Density Function
PFS	Prime Focus Spectroscopy Survey
PT	Perturbation Theory
RPT	Renormalised Perturbation Theory
SPT	Standard Perturbation Theory
WFIRST	Wide Field Infrared Survey Telescope

Symbols and Conventions

β_n	Modal coefficients
$\mathbf{1}_{ij} = \delta^K_{ij}$	Kronecker symbol
$\boldsymbol{k}_{1\ldots n}$	Short-hand notation for the sum $\boldsymbol{k}_1 + \ldots + \boldsymbol{k}_n$
$\boldsymbol{v}(\boldsymbol{x})$	Peculiar velocity at comoving position $\boldsymbol{x}$
$\delta(\boldsymbol{x})$	Matter over-density at comoving position $\boldsymbol{x}$
$\delta_D(\boldsymbol{x})$, $\delta_D(\boldsymbol{k})$	Dirac delta distribution
$\delta_g(\boldsymbol{x})$	Galaxy over-density at comoving position $\boldsymbol{x}$
$\ell(r)$	Line correlation function
$\Gamma^{(n)}$	n-th order multi-point propagator
$\int_{\boldsymbol{k}}$	Short-hand notation for Fourier space integrals, equivalent to $\int \frac{\mathrm{d}^3 k}{(2\pi)^3}$
C	Covariance matrix
Ω_m	Matter density parameter today
Ω_Λ	Dark energy density parameter today
$\overline{n}$	Average number density (usually of galaxies)
Φ	Gravitational potential
σ	Standard deviation
ξ	Two-point correlation function
ζ	Three-point correlation function
A_s	Amplitude of scalar fluctuations
$B(k_1, k_2, k_3)$	Bispectrum
b_i, γ_i	Galaxy bias parameters
c	Speed of light, often set to one
$D(z)$	Linear growth factor as a function of redshift
$D_A(z)$	Angular diameter distance at redshift z
f	Linear growth rate of structures
F_n	Standard perturbation theory kernels for densities
G_n	Standard perturbation theory kernels for velocities
$H(z)$	Hubble rate at redshift z
$iB(k)$	Integrated bispectrum
$K(\boldsymbol{k}_1, \boldsymbol{k}_2)$	Fourier transform of the second-order Galileon operator
$L(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$	Fourier transform of the third-order Galileon operator
n_s	Spectral index
$P(k)$	Power spectrum
$Q(k_1, k_2, k_3)$	Reduced bispectrum
$Q_n(k_1, k_2, k_3)$	Basis functions for modal decomposition
$\mathscr{G}_2, \mathscr{G}_3$	Second- and third-order Galileon operators
$\mathscr{H}(z)$	Comoving Hubble rate at redshift z
$\mathscr{H}_n$	n-th order generalised Wiener-Hermite functional

Chapter 1

Introduction

Cosmology (from the Greek $\kappa\acute{o}\sigma\mu o\varsigma$ for "world" and -$\lambda o\gamma\acute{\iota}\alpha$ for "study of") has greatly benefited from experimental progress achieved in the last century, which has culminated in the formulation of an established model for our Universe — the so-called Λ cold dark matter (ΛCDM) model. This development was initiated by Edwin Hubble in the 1920s who discovered that distant galaxies move faster away from us the farther away they are (Hubble, 1929). Believing that our Milky Way is not a special place in the cosmos, the same should hold true for any other pair of galaxies. Therefore, these motions can be explained if space itself is expanding, which leads us to conclude that all matter must have originated from a single point if we naively time-reverse the trajectories of the galaxies. Such a singularity would have been extremely dense and hot — a state coined as the "big bang" by Fred Hoyle — and implies that some of its heat should still be observable as a left-over black-body radiation today. Confirmation of this picture came in 1964 when Arno Penzias and Robert Wilson discovered a faint, but uniform background noise in their radio telescope (Penzias and Wilson, 1965), which we now call the Cosmic Microwave Background (CMB).

Around the same time when Hubble published his measurements Fritz Zwicky reported that the velocities of galaxies orbiting the centre of the Coma cluster were much higher than what was expected from the visible mass of the cluster (Zwicky, 1933). Based on this observation he inferred the existence of a "dark" matter component, which gained popularity in the 1970s when further experimental evidence was gathered from clusters (Einasto et al., 1974) as well as galaxies (Ostriker et al., 1974; Rubin et al., 1980; Rubin and Ford, 1970), and theoretical arguments suggested that galactic disks would be unstable without a surrounding halo of dark matter (Ostriker and Peebles, 1973). The big bang and dark matter models became fully intertwined after the *Cosmic Background Explorer (COBE)* had mapped the CMB with unprecedented scope and precision between 1989 and 1993. It found tiny anisotropies in the CMB (Bennett et al., 1996; Smoot et al., 1992), which indicate a large amount of cold, invisible matter in the

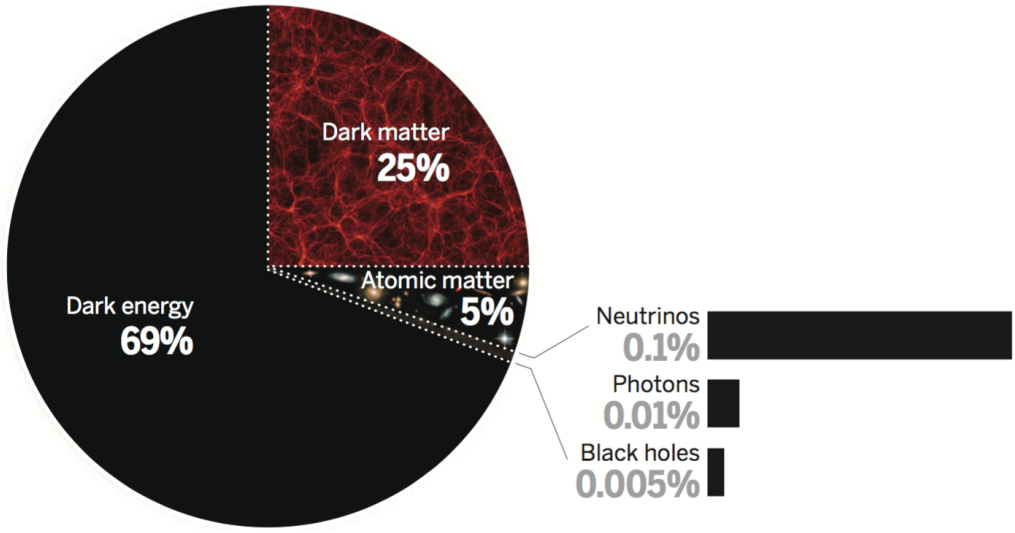


Fig. 1.1 Composition of our Universe into its various forms of energy and their present fractional contribution to the total energy content. The dominant contributions come from dark energy, dark matter and ordinary (baryonic) matter. Less than a per cent of the energy is contained in neutrinos, photons and black holes. Figure credit: (Spergel, 2015)

Universe (hence CDM) is required for these fluctuations to eventually grow into the large-scale distribution of observed galaxies (Peebles, 1982).

The final ingredient, Λ, refers to the cosmological constant, originally introduced by Albert Einstein to counteract any possible expansion of space predicted by his theory of general relativity. However, after Hubble's work he reportedly dismissed his invention, famously calling it "[his] biggest blunder". Interest in the cosmological constant revived in the 1980s, driven by a theoretical prior for geometrically flat universes, together with observational evidence that favoured low matter densities (Gunn and Tinsley, 1975; Peebles, 1984; Turner et al., 1984). A decisive turning point for this debate was the discovery of the accelerated expansion of space, convincing proof of which was presented in 1998 from the observation of distant supernovae (Perlmutter et al., 1999; Riess et al., 1998). This showed that our Universe had to be dominated by some form of energy so far unaccounted for: "dark energy". Curiously, the cosmological constant exhibits exactly the correct properties for this phenomenon, but is not the only explanation.

In the last two decades the ΛCDM model was confronted with measurements from increasingly ambitious experiments, including further CMB missions, such as the *Wilkinson Microwave Anisotropy Probe (WMAP)* (Bennett et al., 2013) and *Planck* (Planck Collaboration et al., 2016a), as well as large-scale structure surveys like the *Two-degree-Field Galaxy Survey* (Colless et al., 2001) and *Baryon Oscillation Spectroscopic Survey (BOSS)* (Dawson et al., 2013). These measurements have given us much more precise estimates of the energy content of our Universe, with the most recent result shown in Fig. 1.1. They demonstrate that over two thirds are indeed taken up by dark

energy, while the matter sector is dominated by dark matter, being five times more abundant than the baryons, and only a negligible contribution is attributed to photons, neutrinos and black holes. Furthermore, it became clear that a six-parameter model was sufficient to match most observations with seemingly remarkable ease, as we will briefly review in Sec. 1.1.

Despite these successes, ΛCDM is a somewhat unsatisfactory model. It requires the existence of new components, but has not given us any clue yet as to what they might be. This is particularly important as it has proven difficult to reconcile these new components with our current understanding of particle physics and general relativity. Moreover, slight inconsistencies amongst different data sets and numerical studies might indicate potential problems (see Sec. 1.2). All these issues are meant to be addressed, at least in parts, by a multitude of forthcoming experiments using a combination of data from supernovae, the CMB, and the large-scale galaxy distribution. This thesis aims to make a contribution to the last of those fields by studying three-point statistics, an analysis method that will — if carefully implemented — allow us to make significantly better use of the data the new experiments will produce. We will lay out our goals as well as the plan for the rest of this thesis in Sec. 1.3.

1.1 ΛCDM and its achievements

The theoretical foundation of the ΛCDM model consists of two assumptions: 1) general relativity is the correct description of gravity, even on the largest distance scales; and 2) our Universe obeys the cosmological principle as proposed by Edward Milne (Milne, 1933, 1934), meaning on sufficiently large scales it appears the same in all directions and from every location. The derivation of the most general metric compatible with these requirements predates Hubble's discovery and was found by various people, which is why it is known as the Friedmann-Lemaître-Robertson-Walker (FLRW) metric (Friedmann, 1922; Lemaître, 1931b; Robertson, 1935; Walker, 1937). It parametrises the expansion of space by the dimensionless scale factor $a(t)$ (normalised to $a(t_0) = 1$ at present time), which is related to the redshifting z of light between its emission at time t and observation today,

$$a(t) = \frac{1}{1+z}. \tag{1.1}$$

The dynamics of the expansion are determined by Einstein's field equations and subject to the energy content and geometry of the universe, the former being characterised by the density ρ and pressure p. This is encapsulated in the Friedmann equations (e.g.

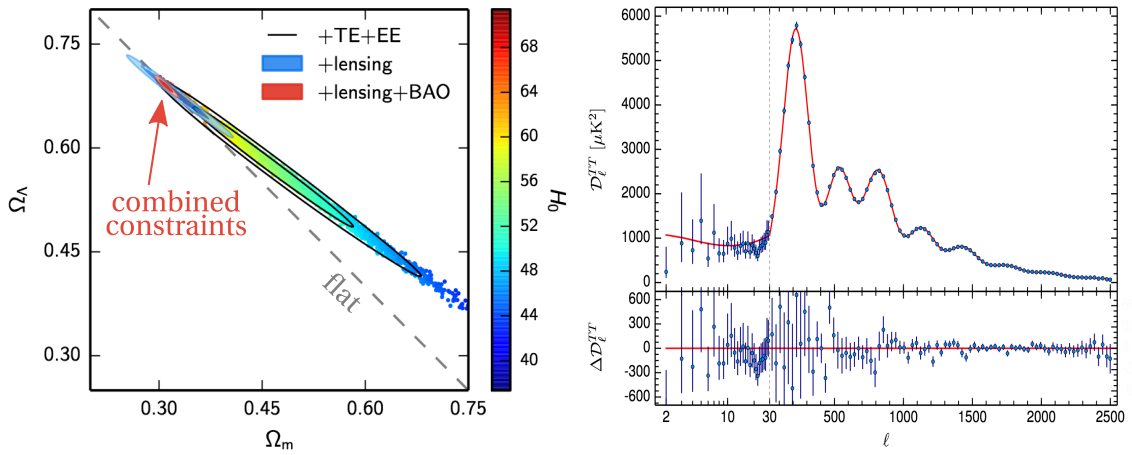


Fig. 1.2 *Left:* Constraints on the matter and cosmological constant density parameters from a combination of *Planck* CMB temperature and polarisation data, CMB lensing and large-scale structure BAO measurements. Samples are colour-coded by the value of H_0, and the dashed line represents a flat ΛCDM model. *Right:* Comparison between the *Planck* CMB temperature power spectrum and the best-fit ΛCDM model; the residuals are shown in the lower panel. Both plots were taken from (Planck Collaboration et al., 2016b).

Weinberg, 2008):

$$H^2(t) = \frac{8\pi G}{3}\rho(t) - \frac{k\,c^2}{a^2(t)}\,, \tag{1.2}$$

$$\frac{\ddot{a}(t)}{a(t)} = -\frac{4\pi G}{3}\left[\rho(t) + 3p(t)\right]. \tag{1.3}$$

The Hubble rate $H(t) \equiv \dot{a}(t)/a(t)$, which enters the first Friedmann equation, defines an important physical scale for any expanding spacetime, as its inverse is a measure for the age and size of the universe. Its value today, usually written as $H_0 \equiv 100\,h$ km/s/Mpc, is a free parameter in ΛCDM and relates the redshift and distance d of nearby galaxies:

$$c\,z = H_0\,d \quad \text{for} \quad z \ll 1\,, \tag{1.4}$$

as established by Hubble. The constant k refers to the curvature, and is positive or negative for closed and open universes, respectively, whereas a flat geometry corresponds to $k = 0$. Following our discussion above, we know that the universe is composed of various forms of energy ρ_X, including ordinary matter (i.e. baryons and electrons) ρ_b, photons ρ_γ, dark matter ρ_c and a cosmological constant ρ_Λ (see also Fig. 1.1). These components are often expressed in terms of their density parameters evaluated at present time,

$$\Omega_X \equiv \frac{\rho_X(t_0)}{\rho_{\rm crit}(t_0)} = \frac{8\pi G\,\rho_X(t_0)}{3\,H_0^2}\,, \tag{1.5}$$

Table 1.1 The six model parameters of ΛCDM and their 68 % error intervals, as determined by *Planck* (Planck Collaboration et al., 2016b, Table 4, 5th column) in 2015. The value of the Hubble constant is given in units of km/s/Mpc.

Parameter description	Symbol	Value
Hubble constant	H_0	67.51 ± 0.64
Physical baryon density parameter	$\Omega_b h^2$	0.02226 ± 0.00016
Physical dark matter density parameter	$\Omega_c h^2$	0.1193 ± 0.0014
Amplitude of fluctuations	$\ln\left(10^{10} A_s\right)$	3.059 ± 0.025
Spectral index	n_s	0.9653 ± 0.0048
Optical depth	τ	0.063 ± 0.014

where ρ_{crit} denotes the energy density the universe acquires in the case it is flat. For vanishing curvature (and neglecting the contributions from photons and neutrinos, which are tiny at present time), the first Friedmann equation implies

$$\Omega_m + \Omega_\Lambda = 1\,, \tag{1.6}$$

where we have combined dark matter and baryons into the single parameter $\Omega_m = \Omega_c + \Omega_b$. Constraints on Ω_m and Ω_Λ are shown in the left-hand panel of Fig. 1.2 from the CMB experiment *Planck* (coloured samples and black contours) in combination with CMB lensing and baryon acoustic oscillation (BAO) measurements from *BOSS* (see Sec. 3.3.3). The CMB alone is degenerate with the value of H_0 in this parameter plane, but together with the other two data sets it clearly favours a flat geometry, in fact: $\Omega_m + \Omega_\Lambda = 1.000 \pm 0.005$ (Planck Collaboration et al., 2016b). For that reason ΛCDM generally assumes a flat universe, such that after the baryon and dark matter density parameters have been specified, Ω_Λ is fixed through Eq. (1.6).

The mean temperature of the CMB, which has been measured by COBE to exquisite precision, $T_{\text{CMB}} = 2.7255 \pm 0.0006$ K (Fixsen, 2009), fixes the photon density parameter Ω_γ. As ΛCDM assumes a standard neutrino sector composed of three distinct species with a combined mass of $\sum m_\nu = 0.06$ eV, the neutrino density can also be computed from T_{CMB} (e.g. Weinberg, 2008). Most other information in the CMB, however, come from the tiny fluctuations in the temperature, which can be condensed into the angular power spectrum, i.e. the strength of fluctuations as a function of inverse angular scale (multipole number ℓ). In order to be able to explain these data, shown in the right-hand panel of Fig. 1.2, ΛCDM includes three more model parameters. They describe the overall amplitude of the fluctuations A_s, and whether fluctuations with longer wavelengths are generally stronger than those with shorter wavelengths, encoded by the spectral index n_s. Finally, once our Universe has produced its first stars, they begin to ionise the surrounding matter, creating free electrons which scatter the CMB photons. This process leads to a damping of the angular power spectrum, which is parametrised

by the optical depth τ. Along with Ω_b, Ω_c and the Hubble rate, it is possible to find a remarkably accurate prediction that is depicted as the red line in Fig. 1.2 and the corresponding parameter values, which have been determined with per cent level precision, are given in Table 1.1. An even bigger success of ΛCDM is that the same set of parameters (without the need for adjusting) also accurately matches statistics derived from other data sets, such as the CMB polarisation, the large-scale distribution of galaxies and weak gravitational lensing (e.g. Aubourg et al., 2015).

1.2 Challenges for ΛCDM

While a cosmological constant fits all our past measurements extremely well, it poses a major theoretical problem. That is because the cosmological constant is equivalent with the energy of the vacuum (Nernst, 1916; Zel'Dovich, 1967; Zel'dovich, 1968), whose theoretical prediction based on quantum field theory gives a vastly different answer than the observed value inferred from $\Omega_\Lambda \approx 0.69$. Depending on the exact details of the quantum field theory computation the discrepancy ranges between 30 and 120 orders of magnitude — a huge number in any case (Burgess, 2013; Weinberg, 1989). A potential solution centres on the argument that the vacuum energy contribution might be cancelled exactly, for instance by some as of yet unknown symmetry mechanism. The accelerated expansion of space can then be realised either by a new dynamical source of energy (Ratra and Peebles, 1988), or by modifying gravity on very large distance scales (e.g. Clifton et al., 2012; Koyama, 2016) and thus giving up one of the foundations of ΛCDM. Both approaches have observational consequences as they alter the equation of state of dark energy, $p_{DE} = w\,\rho_{DE}$, and the growth rate of structures f (for a definition see Sec. 2.2.2). In ΛCDM the equation of state parameter is fixed to $w = -1$, but in more general models it can evolve with time, which is often parametrised as

$$w(a) = w_0 + w_a\,(1-a)\,. \tag{1.7}$$

State-of-the-art constraints from a combination of different experiments are displayed in the left-hand panel of Fig. 1.3, all of which are still in very good agreement with a cosmological constant, indicated by the dashed lines. A similar situation applies to the value of the growth rate, which is predicted to be $f \approx \Omega_m^{5/9}$ in ΛCDM (Bouchet et al., 1995), and no significant deviations from that prediction have been found so far. An overview of recent measurements is given in the right-hand panel of Fig. 1.3, which plots the combination $f\,\sigma_8$ as a function of redshift (σ_8 is related to the amplitude of fluctuations A_s, for a definition see Eq. 3.25). The data points correspond to various large-scale structure surveys, whereas the grey band depicts the ΛCDM value with the allowed uncertainties for Ω_m and σ_8 from *Planck*. The agreement is good over the

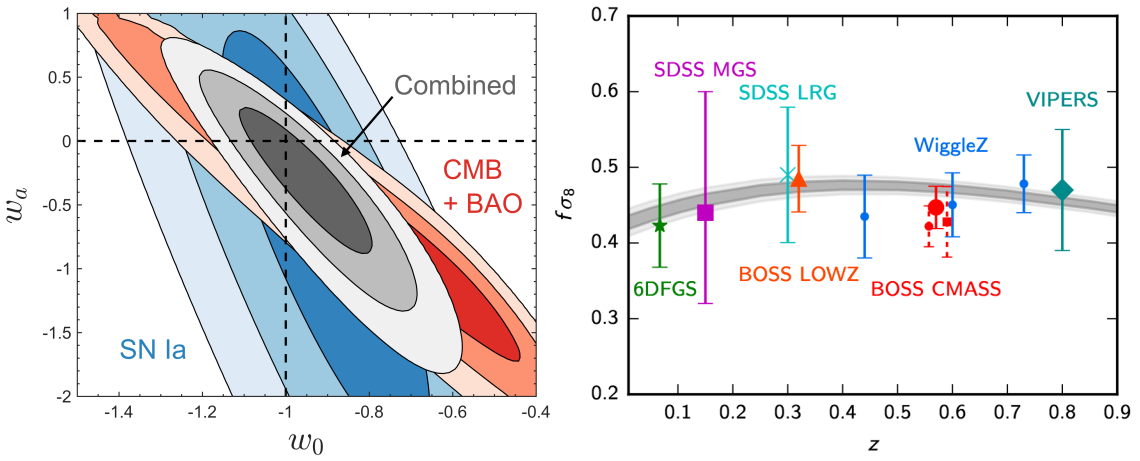


Fig. 1.3 *Left:* constraints on the equation of state parameters of a time-evolving dark energy component (from Huterer and Shafer, 2018). The three contours contain 68 %, 95 % and 99 % of the likelihood and are shown for various most recent data sets: *Planck*, *BOSS* and the JLA supernovae sample. Dashed lines indicate the values of a cosmological constant. *Right:* constraints on the combination of growth rate f and amplitude of fluctuations σ_8 from a variety of large-scale structure surveys (from Planck Collaboration et al., 2016b). The grey band is the ΛCDM prediction for the allowed range of values from *Planck*.

entire redshift range, although Macaulay et al. (2013) have argued that the recent *BOSS* measurements prefer a slightly lower growth rate.

The cosmological constant is not the only challenge for ΛCDM. Until now cold dark matter has evaded all direct detection experiments carried out in laboratories, so it remains unknown what kind of particle could account for it. Moreover, there is a range of observations on scales of galaxies and clusters of galaxies that seem to be in conflict with the dark matter predictions from numerical simulations. These include the "missing satellites problem", which paraphrases the over-abundance of the predicted number of substructures in dark matter halos (Klypin et al., 1999; Moore et al., 1999). It is possible that these substructures lose all of their baryons, which would render them undetectable by optical instruments, but as has been pointed out by Boylan-Kolchin et al. (2012) some of the subhalos are supposed to be too dense to let this occur — the "too big to fail problem". Another discrepancy relates to the density profiles of these halos, which have a pronounced central peak in all ΛCDM models that has been found to be incompatible with observations of dwarf galaxies (Dubinski and Carlberg, 1991; Walker and Peñarrubia, 2011). These and similar problems can possibly be explained after baryons have been properly accounted for in the numerical simulations, but without knowledge of the exact properties of the dark matter particle they might also point towards alternatives, such as warm and self-interacting dark matter, or again modifications of general relativity (for a more comprehensive overview see Bull et al., 2016).

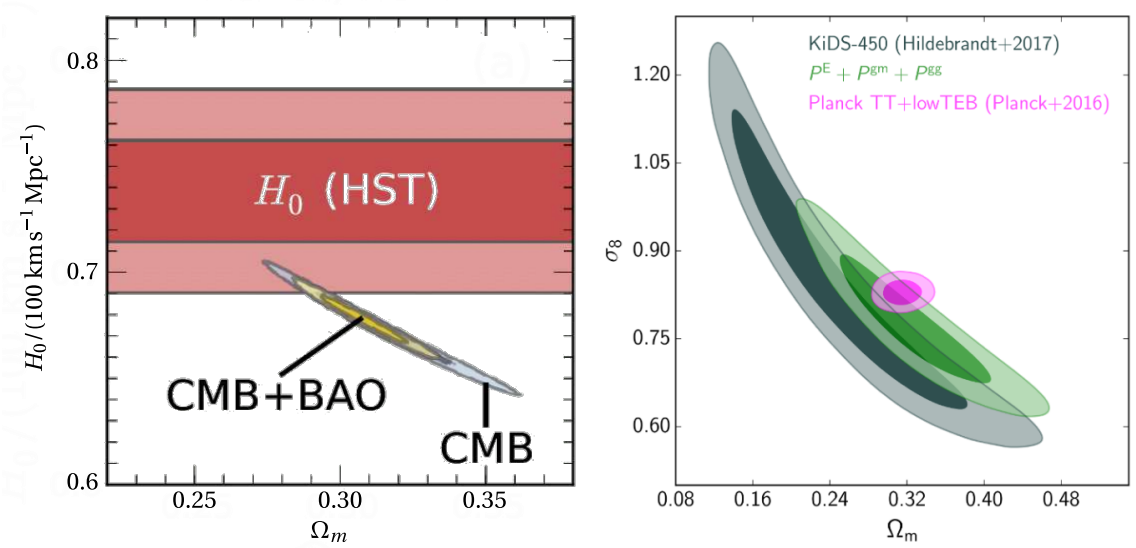


Fig. 1.4 *Left:* Tension between values of the Hubble constant inferred from supernovae experiments (red band) and a combination of CMB and BAO data from *Planck* and *BOSS*, assuming a ΛCDM model (figure from Mortonson et al., 2013). *Right:* Constraints on the matter density parameter and amplitude of fluctuations from weak lensing shear measurements by *KiDS*, and a combination of shear, galaxy lensing and galaxy clustering using *GAMA*, compared to the CMB (figure from van Uitert et al., 2018).

With the advent of more and more precise experiments, slight tensions have arisen between different data sets when adopting the ΛCDM model. The most prominent one features in measurements of the Hubble constant from the local universe (via supernovae) and high-redshift universe (via the CMB and BAO), which are both plotted in the left-hand panel of Fig. 1.4. Compared to the value inferred by *Planck* (see Table 1.1), the former prefers a somewhat higher H_0, and the most recent analysis finds $H_0 = 73.52 \pm 1.62$ km/s/Mpc (Riess et al., 2018), which is about $3.8\,\sigma$ in tension with *Planck*. The credibility of the CMB measurement is enhanced by the independent determination of H_0 from BAOs. While BAOs are not able to constrain the Hubble constant on their own, this is possible once a prior on $\Omega_b\,h^2$ is included, for instance from big bang nucleosynthesis, which yields $H_0 = 67.0 \pm 1.3$ km/s/Mpc (Aubourg et al., 2015). Alternatively, BAOs can be used to calibrate the distance ladder for distant supernovae as opposed to the usual Cepheid variable stars, which also results in a low H_0 value (Aubourg et al., 2015). This might either point towards a potential and unaccounted systematic in the Cepheid calibration, or an extension of the ΛCDM model. Another, but less severe, tension has been reported for the combination of σ_8 and Ω_m from the CMB and weak gravitational lensing. This is shown in the right-hand panel of Fig. 1.4, displaying a $\sim 2\,\sigma$ discrepancy between the shear measurement of the *Kilo Degree Survey (KiDS)* (grey contours) and *Planck* (pink contours). This tension is mostly alleviated when the shear catalogue is combined with the positional data of galaxies from the *Galaxy and Mass Assembly (GAMA)* survey (green contours) (van Uitert et al., 2018) and

a similar result was recently published based on the first year data of the *Dark Energy Survey* (DES Collaboration et al., 2017). In order to resolve the tension in both cases or to discover what is causing them, it will be crucial to further reduce uncertainties in the cosmological parameters.

1.3 Motivation and outline for this thesis

The primary goal of many upcoming large-scale structure surveys, such as the *Dark Energy Spectroscopic Instrument (DESI)* (DESI Collaboration et al., 2016) and *Large Synoptic Survey Telescope (LSST)* (Abell et al., 2009) around 2020, as well as the two space missions, *Euclid* (Amendola et al., 2018) and the *Wide Field Infrared Survey Telescope (WFIRST)* (Spergel et al., 2015) between 2022 and 2025, is to reveal the true nature of dark energy and thus solve one of the major issues in modern physics. As discussed in the last section, achieving this goal requires precise measurements of various cosmological parameters, most importantly the dark energy equation of state and the growth rate.

Shrinking parameter uncertainties critically depends on our ability to make optimal use of the available data, which motivates the main topic of this thesis: the application of three-point statistics, e.g. the bispectrum, to the clustering of galaxies. As such, this work explores an analysis technique that goes beyond the current standard of the field, which is build on the two-point correlation and power spectrum of the galaxy distribution. Both of these two-point statistics have a high signal-to-noise, are relatively easy to measure and we have developed a robust understanding of their theoretical predictions as well as observational systematics. However, unlike for the CMB, two-point statistics of large-scale structures are not sufficient to extract all available information as gravitational evolution inevitably excites a hierarchy of higher-order statistical moments. In addition to the gain of information, three-point correlations distinguish different shapes of structures, for instance spherical from elongated ones. This property allows us to break degeneracies that otherwise afflict two-point statistics and thus significantly reduces parameter uncertainties. That concerns, in particular, the amplitude of fluctuations and growth rate, suggesting that three-point statistics could be integral to the scientific goals of the aforementioned surveys. Finally, by detecting distinctive clustering shapes, three-point statistics also provide a window to test inflation, a hypothesised phase of accelerated expansion in the early universe.

Widespread application of three-point statistics has so far been impeded in part by the availability of large enough data sets, but especially due to increased conceptual challenges compared to its lower-order analogue. These include an inflated size of the covariance matrix, making it difficult to estimate, as well as the theoretical modelling beyond the linear regime, both of which will be addressed in this thesis. We will begin by establishing a theoretical foundation for the evolution of density fluctuations in

Chap. 2 and introduce the concept of galaxy bias. In Chap. 3 we define the standard clustering statistics and illustrate the additional cosmological information contained in three-point statistics. The next two chapters deal with the line correlation function, an alternative three-point measure that specifically targets information not contained in two-point correlations and is therefore possibly able to reduce the dimensionality of the covariance matrix. Chap. 4 first extends the definition of the line correlation function to account for anisotropies and shows that it is sensitive to the growth rate, based on simplified two-dimensional mock data and the Zel'dovich approximation. Subsequently we analyse its covariance properties and the impact of shot noise, supported by a large suite of N-body simulations in Chap. 5. Moreover, we evaluate its sensitivity on a set of cosmological parameters using a Fisher analysis and show that it is particularly adept at constraining galaxy bias and the amplitude of fluctuations. We widen the scope of this analysis in Chap. 6, where we compare various alternative three-point measures, finding that the line correlation function is unlikely to provide a lossless compression of the full three-point covariance matrix, whereas a modal decomposition can shrink its size by at least an order of magnitude. Chap. 7 is dedicated to improving the galaxy bias modelling of the bispectrum in the non-linear regime, in order to make it consistent with comparable state-of-the-art power spectrum models. We develop a new formalism that lets us circumvent the common bias renormalisation problem and demonstrate how the new model extends the range of validity by fitting it to data from a mock galaxy catalogue. Finally, Chap. 8 gives our conclusions and provides an outlook.

Chapter 2

The evolution of density fluctuations

As early as the 1930s and largely supported by the observations of Edwin Hubble (Hubble, 1926, 1934), it began to become clear that galaxies are not distributed at random in our Universe. These initial observations have by now grown into extensive surveys of the sky, dedicated to mapping the positions of millions of galaxies. An example of such a survey extending to about a redshift of 0.2 is shown in Fig. 2.1 and reveals an intriguing structure of our Universe on scales of several Mpc and beyond. It suggests that galaxies reside in dense clusters, that are connected by large filaments and sheets, which ultimately surround vast regions of seemingly empty space.

The simple existence of these structures begs two immediate questions. First, where do they come from? And second, how do they form and evolve? According to our current knowledge, these questions find their resolution in a cosmic tale that spans almost 14 billion years and starts out in the very first moments of our Universe. At that time a phase of rapid accelerated expansion, appropriately referred to as inflation, stretches tiny patches of spacetime and their surrounding quantum fluctuations into macroscopic proportions. These fluctuations provide the primordial inhomogeneities in the matter and radiation field that can subsequently grow through gravitational amplification into more complicated structures, as was originally envisioned by George Lemaître (Lemaître, 1931a; Lemaitre, 1934). However, initially gravity is in contest with pressure gradients and a still quickly expanding Universe that counteract the growth of the primordial fluctuations. Only when the Universe has cooled sufficiently it is possible to form compact structures, such as dark matter halos, galaxies, and eventually, our Solar System.

The aim of this chapter is to give an overview of various stages that are relevant for describing the primordial quantum fluctuation's evolution into galaxies. We will place particular emphasis on the main theoretical frameworks that allow us to accurately describe the non-linear large-scale matter field and its relation to galaxies, which is required to interpret the data gathered by current and future galaxy redshift surveys.

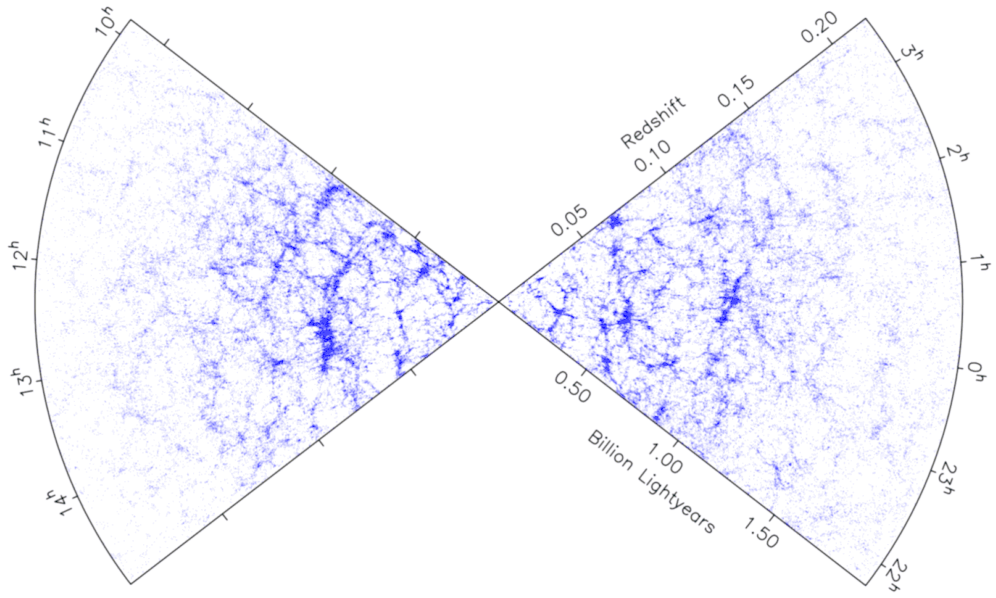


Fig. 2.1 Distribution of galaxies recorded by the 2dF galaxy redshift survey. The picture shows a 3° slice in declination and contains a total of 62559 galaxies out to a redshifts of ~ 0.2. (From Colless et al., 2003).

2.1 Initial conditions

Inflation was originally introduced by (Albrecht and Steinhardt, 1982; Guth, 1981; Linde, 1982), partly to address the horizon problem of cosmology, which paraphrases the startling observation that regions that cannot have been in causal contact with one another display the same CMB temperature. Only afterwards it was realised that inflation at the same time serves as a method for generating primordial density fluctuations (Bardeen et al., 1983; Guth and Pi, 1982; Hawking, 1982; Starobinsky, 1982). Although substantial progress has been made, the exact details of inflation remain unknown until today, which is why we will illustrate the generation of initial conditions for the simplest possible incarnation — a single scalar field that is slowly evolving in its potential.

2.1.1 Single scalar-field inflation

If we imagine the universe contains a single form of energy, what properties does it need to have in order to drive a period of inflation, that is, a period where the universe undergoes accelerated expansion? For an homogeneous and isotropic universe, these properties are characterised by the density $\rho(t)$ and pressure $p(t)$, which are directly related to the behaviour of the scale factor $a(t)$ by the Friedmann equations (1.2) and (1.3). The difference between a decelerating ($\ddot{a} < 0$) and accelerating ($\ddot{a} > 0$) expansion

becomes particularly clear in terms of the comoving Hubble horizon, $\mathscr{H}^{-1} \equiv (a\,H)^{-1}$, which is a strictly increasing function of time in the former case, but decreasing in the latter. From Eq. (1.3) we see that this requires the energy content to exert a negative pressure, $p < -\rho/3$, which neither ordinary matter nor radiation can account for.

On the other hand, the energy density and pressure of a homogeneous scalar field $\varphi(t)$ that is subject to the potential $V(\varphi)$ can be shown to be

$$\rho_\varphi \simeq \frac{1}{2}\dot{\varphi}^2 + V(\varphi)\,, \quad p_\varphi \simeq \frac{1}{2}\dot{\varphi}^2 - V(\varphi)\,. \tag{2.1}$$

If the field, which we will now refer to as the inflaton, is furthermore in a configuration in which the potential energy dominates over its kinetic energy, we get $p_\varphi \simeq -\rho_\varphi$, satisfying the bound above. However, in order to maintain this field configuration for an extended amount of time, the inflaton must also be evolving slowly. According to its equation of motion,

$$\ddot{\varphi} + 3H\dot{\varphi} + V'(\varphi) = 0\,, \tag{2.2}$$

this implies that $|\ddot{\varphi}| \ll 3|H\dot{\varphi}|, |V'(\varphi)|$, and as it turns out, all of these conditions can be summarised by the two slow-roll parameters

$$\epsilon(\varphi) \equiv \frac{1}{16\pi G}\left[\frac{V'(\varphi)}{V(\varphi)}\right]^2\,, \quad \eta(\varphi) \equiv \frac{1}{8\pi G}\frac{V''(\varphi)}{V(\varphi)}\,, \tag{2.3}$$

which guarantee that $a(t) \sim \exp(H\,t)$, as long as both of them are small. The inflationary potential must therefore be either very flat, or alternatively, take rather large values.

Once the slow-roll conditions are violated, inflation ends and is followed by a phase called reheating. During this phase the inflaton eventually decays into particles of the standard model and the time when these particles have thermalised is usually taken to be the beginning of the "hot big bang". The process of reheating depends on the exact couplings between the inflaton and other particles and thus likely relies on a great deal of unknown physics. Fortunately, and as we will see in the next section, this has no impact on the primordial fluctuations that we are interested in.

2.1.2 Scalar perturbations from inflation

Let us now consider how tiny deviations from a homogeneous inflaton field give rise to the primordial density fluctuations. Writing the position dependent scalar field as $\varphi(\boldsymbol{x}, t) = \bar{\varphi}(t) + \delta\varphi(\boldsymbol{x}, t)$, we can derive an equation of motion for the inflaton perturbations that is valid in the slow-roll regime. For a specific Fourier mode $\boldsymbol{k}$ it is given

by

$$\ddot{\delta\varphi}(\boldsymbol{k}) + 3H\dot{\delta\varphi}(\boldsymbol{k}) + \frac{k^2}{a^2}\delta\varphi(\boldsymbol{k}) = 0\,, \tag{2.4}$$

which corresponds to an harmonic oscillator with scale dependent frequency k/a and a damping term due to the Hubble expansion. This analogy allows us to promote the perturbations $\delta\varphi_{\boldsymbol{k}}$ to quantum operators,

$$\delta\varphi(\boldsymbol{k},t) \longrightarrow \widehat{\delta\varphi}(\boldsymbol{k},t) = v(\boldsymbol{k},t)\,\hat{a}(\boldsymbol{k}) + v^*(-\boldsymbol{k},t)\,\hat{a}^\dagger(-\boldsymbol{k})\,, \tag{2.5}$$

where $v_{\boldsymbol{k}}$ denotes a complex mode function and the creation and annihilation operators satisfy the same commutation relation $\left[\hat{a}(\boldsymbol{k}),\,\hat{a}^\dagger(\boldsymbol{k}')\right] = (2\pi)^3\,\delta_D(\boldsymbol{k}-\boldsymbol{k}')$ as in the usual harmonic oscillator case. The normalisation of the mode functions, together with the selection of an appropriate vacuum state provide the necessary boundary conditions to find a solution for Eq. (2.4). In the limit $\epsilon(\varphi) \to 0$ this leads to the following vacuum fluctuations on super-horizon scales, i.e. $k \ll a\,H$:

$$\langle\widehat{\delta\varphi}(\boldsymbol{k},t)\,\widehat{\delta\varphi}(\boldsymbol{k}',t)\rangle \equiv (2\pi)^3\,P_{\delta\varphi}(k)\,\delta_D(\boldsymbol{k}+\boldsymbol{k}') = (2\pi)^3\,\delta_D(\boldsymbol{k}+\boldsymbol{k}')\,\frac{H^2}{2k^3}\,, \tag{2.6}$$

where we have defined the power spectrum of inflaton perturbations as $P_{\delta\varphi}(k) = H^2/2k^3$. The dimensionless power spectrum $\Delta^2_{\delta\varphi}(k) \equiv k^3\,P_{\delta\varphi}(k)/2\pi^2$ is therefore independent of scale — amazingly, this prediction was already made by Harrison and Zel'dovich based purely on naturalness arguments prior to the invention of inflation. In inflation, however, there are slight deviations from scale-invariance due to the time dependence of the Hubble rate during inflation (i.e. $\epsilon(\varphi)$ is not exactly zero). This is parametrised by the spectral index n_s, which can be related to the slow-roll parameters

$$n_s - 1 \equiv \frac{\mathrm{d}\ln\Delta^2_\varphi}{\mathrm{d}\ln k} \approx 2\eta(\varphi) - 6\epsilon(\varphi)\,. \tag{2.7}$$

Generally this gives spectral indices that are slightly less than one, which is a testable prediction of inflation and has been well confirmed by recent CMB measurements of *Planck*. Another important prediction is that the perturbation spectrum is Gaussian, which is a consequence of the linearity of Eq. (2.4). That entails that the expectation values of three or more fluctuation modes $\widehat{\delta\varphi}(\boldsymbol{k})$ either vanish or are expressible through the power spectrum[1]. Deviations from Gaussianity in the single-field case are possible for certain inflaton potentials, but are thought to be rather small (Acquaviva et al., 2003; Maldacena, 2003).

[1] A more precise definition of a Gaussian field is given in Sec. 3.1.1.

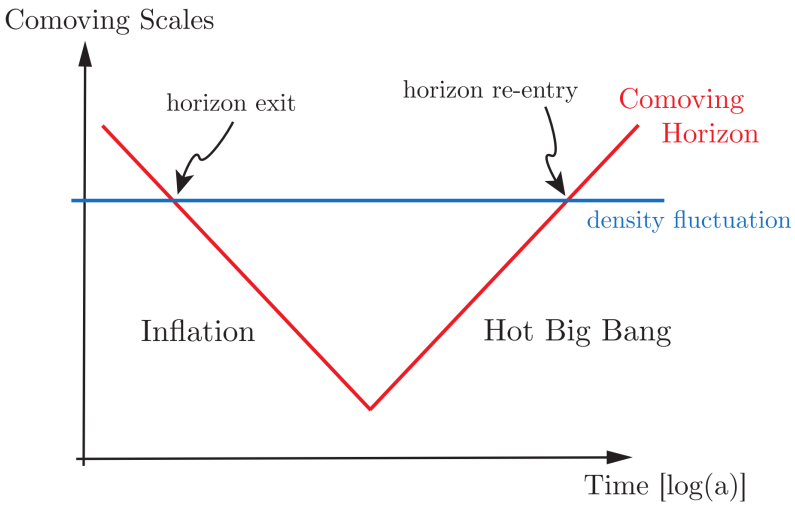


Fig. 2.2 Illustration of the comoving horizon (depicted by the red line) and a comoving scale (in blue) as a function of scale factor (from Baumann, 2009). During inflation the comoving horizon is shrinking, whereas it expands afterwards. That causes density fluctuations on certain comoving scales, which are time-independent, to leave the horizon during inflation. They are preserved until they eventually re-enter the horizon.

Eventually the slow-roll approximation is going to break down, which would invalidate our treatment so far. However, before the inflaton dynamics become helplessly complicated we can relate its vacuum fluctuations to perturbations in the spacetime metric, which we have ignored so far. By doing so we have implicitly made a particular choice of gauge, where the spatial part of the metric perturbations, represented by the potential $\widehat{\Psi}(\boldsymbol{k})$, vanishes. It is possible to define gauge-invariant combinations of metric and inflaton perturbations, such as the comoving curvature perturbation[2], which during inflation is given by (Baumann, 2009)

$$\widehat{\mathscr{R}}(\boldsymbol{k}) = \widehat{\Psi}(\boldsymbol{k}) + \frac{H}{\dot{\varphi}}\widehat{\delta\varphi}(\boldsymbol{k})\,, \tag{2.8}$$

and has the remarkable property that it is conserved on super-horizon scales. Using this fact we gain the following qualitative picture (see Fig. 2.2): First, inflation generates fluctuations on all scales k^{-1}, which eventually leave the horizon as $(a\,H)^{-1}$ (depicted by the red line) is shrinking during inflation. At the time of horizon crossing, i.e. $k = a(t_\star)\,H(t_\star)$, the fluctuations freeze out and fix the corresponding spectrum of comoving curvature perturbations — using Eq. (2.8) we get:

$$P_{\mathscr{R}}(k) = \left(\frac{H_\star}{\dot{\varphi}_\star}\right)^2 P_{\delta\varphi}(k)\Big|_{a_\star H_\star = k}\,. \tag{2.9}$$

[2]The comoving curvature perturbation measures the intrinsic spatial curvature of constant-φ hypersurfaces.

Moreover, on super-horizon scales the quantum nature of all three operators in Eq. (2.8) disappears (Guth and Pi, 1985) and their vacuum expectation values can be regarded as the ensemble averages of classical stochastic fields.

After inflation has ended the comoving horizon starts growing again and upon re-entry we can use the now classical field $\mathscr{R}(\boldsymbol{k})$ to express the fluctuations completely in terms of Ψ, or equivalently the gravitational potential Φ (Dodelson, 2011). Once inside the horizon, the gravitational potential evolves in conjunction with the various species in the universe — dark matter, baryons and radiation. If a mode enters during radiation domination, the main part of the energy content undergoes acoustic oscillations because radiation pressure counteracts the gravitationally driven growth of perturbations. This causes the gravitational potential to decay with scale factor, while during matter domination it stays constant and thus allows for efficient growth as we will see in the next section. All these effects on the potential, from horizon entry until the point where it is time independent again, are usually encoded in the transfer function $T(k)$. At a time where $a = a_T$ the gravitational potential is related to the comoving curvature perturbation from inflation as follows (Dodelson, 2011)

$$\Phi(\boldsymbol{k}, a_T) = -\frac{3}{5}\mathscr{R}(\boldsymbol{k})\, T(k)\,, \tag{2.10}$$

where the exact definition of a_T is arbitrary as long as it corresponds to a time where $\Phi \approx$ const., i.e. after recombination but well before dark energy comes to dominate the universe's expansion.

2.2 Large-scale matter perturbations

We have seen how quantum fluctuations in the inflaton field get converted into classical perturbations in the gravitational potential. In this section we consider how these perturbations get in turn imprinted onto the dark matter field during the phase of matter domination[3] , where we can safely ignore the impact from baryons and radiation. We will focus, in particular, on the large-scale dark matter field, so that we can study its evolution perturbatively. In addition, we restrict ourselves to modes well inside the horizon, which warrants a Newtonian treatment.

2.2.1 The fluid equations

We consider dark matter as a smooth, pressure-less fluid that is characterised by its density $\rho(\boldsymbol{x}, \tau)$ and peculiar velocity field $\boldsymbol{v}(\boldsymbol{x}, \tau)$, which we write as functions of the

[3] During radiation domination the growth of dark matter perturbations is suppressed and only increases logarithmically with the scale factor due to the decaying gravitational potential, known as the Mészáros effect (Mészáros, 1974). This effect is already incorporated in the transfer function.

comoving position $\boldsymbol{x}$ and conformal time $\tau = \int \mathrm{d}t/a(t)$. At early times in the matter dominated epoch, or equivalently on large enough scales, dark matter flows are coherent, so that the exact particle dynamics are well approximated by neglecting any dispersion in the velocity field. This approximation breaks down when different flows begin to interact with each other in an event called "shell crossing", which eventually leads to the formation of compact objects (see Sec. 2.3). In the present case, however, the behaviour of the dark matter can be well described by the continuity and Euler equations

$$\frac{\partial \delta(\boldsymbol{x},\tau)}{\partial \tau} + \nabla \cdot \Big[\Big(1+\delta(\boldsymbol{x},\tau)\Big)\, \boldsymbol{v}(\boldsymbol{x},\tau) \Big] = 0\,, \tag{2.11}$$

$$\frac{\partial \boldsymbol{v}(\boldsymbol{x},\tau)}{\partial \tau} + \mathscr{H}(\tau)\, \boldsymbol{v}(\boldsymbol{x},\tau) + \Big[\boldsymbol{v}(\boldsymbol{x},\tau)\cdot\nabla \Big]\, \boldsymbol{v}(\boldsymbol{x},\tau) = -\nabla \Phi(\boldsymbol{x},\tau)\,, \tag{2.12}$$

where we have decomposed the density into a background value $\overline{\rho}$ and perturbation δ,

$$\rho(\boldsymbol{x},\tau) \equiv \overline{\rho}(\tau)\Big[1+\delta(\boldsymbol{x},\tau)\Big]\,. \tag{2.13}$$

The evolution of the background density is determined by the Friedmann equations (1.2) and (1.3) can be expressed in terms of the dark matter density parameter and comoving Hubble rate as

$$\overline{\rho}(\tau) = \frac{3\Omega_m}{8\pi G}\left(\frac{\mathscr{H}(\tau)}{a(\tau)}\right)^2 \propto \frac{1}{a(\tau)^3}\,. \tag{2.14}$$

The dark matter perturbations, on the other hand, couple to the gravitational potential via the Poisson equation and using Eq. (2.14) we have

$$\nabla^2 \Phi(\boldsymbol{x},\tau) = \frac{3}{2}\,\Omega_m\, \mathscr{H}(\tau)^2\, \delta(\boldsymbol{x},\tau)\,, \tag{2.15}$$

which fixes the fluid dynamics completely. An important consequence of neglecting effects stemming from the velocity dispersion is that the only source term in Eq. (2.12) is the gradient of the gravitational potential. This source alone is not capable of inducing curl modes in the velocity field and any initial ones will quickly become subdominant (Bernardeau et al., 2002). Until shell crossing occurs it is therefore an equally good approximation to take the velocity field as purely potential, which has been demonstrated explicitly by Pichon and Bernardeau (1999); Pueblas and Scoccimarro (2009) and implies we can work with the velocity divergence $\theta(\boldsymbol{x},\tau) = \nabla\cdot\boldsymbol{v}(\boldsymbol{x},\tau)$ instead.

2.2.2 Linear growth

Our goal in the next sections will be to solve the system of fluid equations perturbatively. Let us start by determining the linear solution, in which case we can neglect all terms

that involve the products of fields. Taking the divergence of Eq. (2.12) and using Eq. (2.11) and (2.15) to eliminate the velocity divergence, we obtain a second-order differential equation for the density perturbation:

$$\frac{\partial^2 \delta(\boldsymbol{x},\tau)}{\partial \tau^2} + \mathscr{H}(\tau)\,\frac{\partial \delta(\boldsymbol{x},\tau)}{\partial \tau} - \frac{3}{2}\,\Omega_m\,\mathscr{H}(\tau)^2\,\delta(\boldsymbol{x},\tau) = 0\,. \tag{2.16}$$

This expression reveals that the growth of structures, at least at linear level, is independent of scale, because if we were to Fourier transform Eq. (2.16) all modes $\delta(\boldsymbol{k})$ would grow at the same rate. This means the solutions can be factorised into a spatial and time-dependent part, and as Eq. (2.16) is second order in the time derivatives, we can expect two independent solutions. In general, we will thus have a superposition of the form

$$\delta(\boldsymbol{x},\tau) = D_+(\tau)\,\Delta_+(\boldsymbol{x}) + D_-(\tau)\,\Delta_-(\boldsymbol{x})\,, \tag{2.17}$$

which can be solved exactly for an Einstein de-Sitter background (a matter dominated universe with $\Omega_m = 1$ and zero curvature), giving the following linear growth factors $D_\pm(\tau)$:

$$D_+(\tau) = a(\tau)\,, \quad D_-(\tau) = a(\tau)^{-3/2}\,. \tag{2.18}$$

The dependence on the scale factor shows that the second solution decays very quickly with the expansion of the universe, while the first one is growing and accordingly the dominant mode for structure formation. From applying the linearised continuity equation we get a corresponding expression for θ:

$$-\frac{\theta(\boldsymbol{x},\tau)}{\mathscr{H}} = f_+\,D_+(\tau)\,\Delta_+(\boldsymbol{x}) + f_-\,D_-(\tau)\,\Delta_-(\boldsymbol{x})\,, \tag{2.19}$$

where we have introduced the logarithmic growth rates $f_\pm \equiv \mathrm{d}\ln D_\pm / \mathrm{d}\ln a$.

The spatial part of the growing mode solution, $\delta_0 \equiv \delta_+$, can be related to the value of the gravitational potential during matter domination from Eq. (2.10). Applying the Poisson equation in Fourier space gives:

$$\delta_0(\boldsymbol{k}) = \frac{2}{5}\,\frac{\mathscr{R}(\boldsymbol{k})\,T(k)}{\Omega_m\,H_0^2}\,k^2\,, \tag{2.20}$$

where we have normalised by the present day value of the Hubble rate, H_0. The spatial fluctuations of the linear density modes are thus characterised by the comoving

curvature perturbations (Eq. 2.9), and allowing for a non-zero spectral index we get

$$P_0(k) \equiv A_s\, k^{n_s}\, T(k)^2 \equiv \frac{4}{50}\left(\frac{H_\star}{H_0}\right)^4 \frac{k^{n_s}}{\Omega_m^2\, \dot{\bar{\varphi}}_\star^2}\, T(k)^2\,. \tag{2.21}$$

This defines the power spectrum amplitude A_s, which can for instance be calibrated against measurements from CMB experiments.

Another way of arriving at the linear theory results is to write the fluid equations in terms of the doublet[4]

$$\Psi_a(\boldsymbol{x},\eta) \equiv \left(\delta(\boldsymbol{x},\eta), -\frac{\theta(\boldsymbol{x},\eta)}{\mathscr{H} f}\right), \tag{2.22}$$

and in combination with the new time variable $\eta = \ln D$ this leads to

$$\frac{\partial}{\partial\eta}\Psi_a(\boldsymbol{x},\eta) + \Omega_{ab}(\eta)\,\Psi_b(\boldsymbol{x},\eta) = 0\,, \quad \text{with} \quad \Omega_{ab}(\eta) = \begin{bmatrix} 0 & -1 \\ -\frac{3}{2}\frac{\Omega_m}{f^2} & \frac{3}{2}\frac{\Omega_m}{f^2} - 1 \end{bmatrix}. \tag{2.23}$$

The general solution to this equation is given by the linear propagator $g_{ab}(\eta,\eta_0)$, which relates the values of the density and velocity divergence at time η to their initial values at η_0, i.e. $\Psi_a(\boldsymbol{x},\eta) = g_{ab}(\eta,\eta_0)\,\Psi_b(\boldsymbol{x},\eta_0)$. In the Einstein de-Sitter case we have $\Omega_m/f^2 = 1$ and the propagator is again a superposition of a growing and decaying mode:

$$g_{ab}(\eta,\eta_0) = \frac{\mathrm{e}^{\eta-\eta_0}}{5}\begin{bmatrix} 3 & 2 \\ 3 & 2 \end{bmatrix} + \frac{\mathrm{e}^{-\frac{3}{2}(\eta-\eta_0)}}{5}\begin{bmatrix} 2 & -2 \\ -3 & 3 \end{bmatrix}. \tag{2.24}$$

While the linear growth factors differ when changing the background model (see Hamilton, 2001, for explicit expressions), this propagator solution is remarkably robust under changes of the cosmology. That is because in most models of interest to us, $\Omega_m/f^2 \approx 1$ to very good accuracy. Most importantly, $f \approx \Omega_m^{5/9}$ (Bouchet et al., 1995) in models with a non-zero cosmological constant and vanishing curvature ($\Omega_m + \Omega_\Lambda = 1$). This simple observation will have important consequences for the perturbation theory kernels, that we will consider in the next section.

2.2.3 Nonlinear Eulerian perturbation theory

Since the perturbations in the dark matter density and velocity are growing, we will encounter a regime where we are no longer able to neglect the terms involving products of these fields. Once they start influencing the dynamics of the fluid, they lead to a phenomenon called "mode coupling", where different scales of the fields do not evolve independently from one another anymore. In this case it is most convenient to study

[4] From now on we will simplify the notation by defining $D \equiv D_+$ and $f \equiv f_+$.

the fluid equations in Fourier space and, upon taking the divergence of Eq. (2.12) and using the notation introduced at the end of Sec. 2.2.2, they read:

$$\frac{\partial}{\partial\eta}\Psi_a(\boldsymbol{k}) + \Omega_{ab}\,\Psi_b(\boldsymbol{k}) = (2\pi)^3 \int\limits_{\boldsymbol{k}_1,\boldsymbol{k}_2} \delta_D(\boldsymbol{k}-\boldsymbol{k}_{12})\,\gamma_{abc}(\boldsymbol{k}_1,\boldsymbol{k}_2)\,\Psi_b(\boldsymbol{k}_1)\,\Psi_c(\boldsymbol{k}_2)\,, \tag{2.25}$$

where we have suppressed the time dependence of all fields. The fundamental mode coupling kernels are given by

$$\gamma_{121}(\boldsymbol{k}_1,\boldsymbol{k}_2) \equiv \alpha(\boldsymbol{k}_1,\boldsymbol{k}_2) = \frac{\boldsymbol{k}_{12}\cdot\boldsymbol{k}_1}{k_1^2}\,, \tag{2.26}$$

$$\gamma_{222}(\boldsymbol{k}_1,\boldsymbol{k}_2) \equiv \beta(\boldsymbol{k}_1,\boldsymbol{k}_2) = \frac{k_{12}^2}{2k_1^2\,k_2^2}\,\boldsymbol{k}_1\cdot\boldsymbol{k}_2\,, \tag{2.27}$$

and all other components of γ_{abc} vanish. They indicate that the evolution of the fields at a certain scale $\boldsymbol{k}$ now also depends on the field values of all pairs of modes $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ whose sum $\boldsymbol{k}_{12}$ is equal to $\boldsymbol{k}$. As shown in detail in Scoccimarro (2001) Eq. (2.25) permits an integral solution over the linear propagator that was already obtained above:

$$\begin{aligned}\Psi_a(\boldsymbol{k},\eta) = g_{ab}(\eta,\eta_0)\,\Psi_b(\boldsymbol{k},\eta_0) + (2\pi)^3 \int_{\eta_0}^{\eta} \mathrm{d}\eta'\, g_{ab}(\eta,\eta') \int\limits_{\boldsymbol{k}_1,\boldsymbol{k}_2} \delta_D(\boldsymbol{k}-\boldsymbol{k}_{12})\,\gamma_{bcd}(\boldsymbol{k}_1,\boldsymbol{k}_2) \\ \times\, \Psi_c(\boldsymbol{k}_1,\eta')\,\Psi_d(\boldsymbol{k}_2,\eta')\,.\end{aligned} \tag{2.28}$$

The idea of Eulerian perturbation theory is to expand Ψ_a in a series of increasing powers of the primordial density fluctuations $\delta_0(\boldsymbol{k})$, which allows us to solve Eq. (2.28) order by order. At n-th order our ansatz reads

$$\Psi_a^{(n)}(\boldsymbol{k},\eta) = (2\pi)^3 \int\limits_{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n} \delta_D(\boldsymbol{k}-\boldsymbol{k}_{1\ldots n})\,\mathscr{F}_a^{(n)}(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n;\eta)\,\delta_0(\boldsymbol{k}_1)\cdots\delta_0(\boldsymbol{k}_n)\,, \tag{2.29}$$

and the functions $\mathscr{F}^{(n)}$, which account for the coupling between different density modes, can in principle be time dependent. Plugging this ansatz into Eq. (2.28) and grouping terms of the same order (same power of primordial density fluctuations) together, yields a recursion relation that fixes the mode coupling kernels. An important conclusion that can be drawn from these recursion relations and the particular Einstein de-Sitter propagator given in Eq. (2.24) is that time and scale dependence of the kernels completely separate. That means what we observed to be the case at linear order is

actually true at all orders of perturbation theory and thus Eq. (2.29) becomes

$$\Psi_a^{(n)}(\boldsymbol{k},\tau) = (2\pi)^3\, D(\tau)^n \int_{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n} \delta_D(\boldsymbol{k}-\boldsymbol{k}_{1\ldots n}) \left[\begin{array}{c} F_n(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n) \\ G_n(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n) \end{array}\right]_a \delta_0(\boldsymbol{k}_1)\cdots\delta_0(\boldsymbol{k}_n)\,, \tag{2.30}$$

(see Appendix B.3 in Scoccimarro et al., 1998b) with the usual kernels, F_n and G_n, for Eulerian perturbation theory. Their recursion relations are far simpler and can be expressed in terms of the two fundamental mode coupling kernels as derived in Goroff et al. (1986); Jain and Bertschinger (1994):

$$\begin{aligned} F_n(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n) = \sum_{m=1}^{n-1} \frac{G_m(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_m)}{(2n+3)\,(n-1)} \Big[&(2n+1)\,\alpha(\boldsymbol{K}_1,\boldsymbol{K}_2)\, F_{n-m}(\boldsymbol{k}_{m+1},\ldots,\boldsymbol{k}_n) \\ &+ 2\beta(\boldsymbol{K}_1,\boldsymbol{K}_2)\, G_{n-m}(\boldsymbol{k}_{m+1},\ldots,\boldsymbol{k}_n)\Big]\,, \end{aligned} \tag{2.31}$$

$$\begin{aligned} G_n(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n) = \sum_{m=1}^{n-1} \frac{G_m(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_m)}{(2n+3)\,(n-1)} \Big[&3\alpha(\boldsymbol{K}_1,\boldsymbol{K}_2)\, F_{n-m}(\boldsymbol{k}_{m+1},\ldots,\boldsymbol{k}_n) \\ &+ 2n\,\beta(\boldsymbol{K}_1,\boldsymbol{K}_2)\, G_{n-m}(\boldsymbol{k}_{m+1},\ldots,\boldsymbol{k}_n)\Big]\,, \end{aligned} \tag{2.32}$$

where $\boldsymbol{K}_1 \equiv \boldsymbol{k}_1+\ldots+\boldsymbol{k}_m$, $\boldsymbol{K}_2 \equiv \boldsymbol{k}_{m+1}+\ldots+\boldsymbol{k}_n$ and $F_1 = 1 = G_1$. Furthermore, as we argued before, the linear propagator is mostly cosmology independent provided that $\Omega_m/f^2 \approx 1$, from which immediately follows that Eq. (2.30) should hold for any background model that is of interest here. All cosmology dependence, such as the dependence on Ω_m and Ω_Λ, is therefore simply contained in the prefactors that involve n powers of the linear growth factor, while the kernels (Eqs. 2.31 and 2.32) are mostly model independent.

A quantity that will appear at various points later in this thesis is the second-order kernel of the dark matter density. After symmetrizing over its two arguments we obtain from Eq. (2.31):

$$F_2(\boldsymbol{k}_1,\boldsymbol{k}_2) = \frac{5}{7} + \frac{\mu}{2}\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right) + \frac{2}{7}\mu^2 = \frac{17}{21} + \frac{\mu}{2}\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right) + \frac{2}{7}\left(\mu^2-\frac{1}{3}\right), \tag{2.33}$$

where $\mu \equiv \boldsymbol{k}_1\cdot\boldsymbol{k}_2/k_1\,k_2$, and on the right hand side we have explicitly written it in terms of a monopole, dipole and quadrupole contribution. These terms can be interpreted as follows: the monopole is simply due to the collapse of a spherically symmetric over-density (in fact, we will encounter the same coefficient in Sec. 2.3.1), while the dipole corresponds to a bulk flow of matter, i.e. it is proportional to $\boldsymbol{v}\cdot\nabla\delta$. Finally, the quadrupole encodes the effect from tidal gravitational forces that distort the shape of the over-densities. These effects are also the main ingredients for all higher orders of perturbation theory, even though more complicated terms such as bulk flows of the tidal field etc. appear.

2.2.4 Lagrangian perturbation theory

An alternative way of describing the emerging non-linearities in the perturbations was proposed by Yakov Zel'dovich (Zel'dovich, 2009). His idea was to trace the the motion of individual fluid cells opposed to evaluating the overall density and velocity field at fixed times. The central quantity of this approach, now commonly called Lagrangian perturbation theory, is the displacement field $\boldsymbol{\Psi}(\boldsymbol{q},\tau)$ (not to be confused with the doublet Ψ_a from before), which relates the final, Eulerian, positions of the fluid elements to their initial ones $\boldsymbol{q}$:

$$\boldsymbol{x}(\boldsymbol{q},\tau) = \boldsymbol{q} + \boldsymbol{\Psi}(\boldsymbol{q},\tau)\,. \tag{2.34}$$

The "initial" coordinate set refers to a time where the universe was completely homogeneous[5], so that the conservation of mass contained in an infinitesimal fluid element, $[1+\delta(\boldsymbol{x},\tau)]\,\mathrm{d}^3x = \mathrm{d}^3q$, implies a relation between the Jacobian of this transformation and the density perturbations,

$$J(\boldsymbol{q},\tau) \equiv \det\left[\delta^K_{ij} + \nabla_{q,j}\,\Psi_i\right] = \frac{1}{1+\delta(\boldsymbol{x},\tau)}\,. \tag{2.35}$$

It is interesting to note that this relation is intrinsically nonlinear, which means that even a small displacement of the fluid cells already carries a large amount of nonlinear information (Scoccimarro and Frieman, 1996). This can be seen more explicitly by writing the density perturbation in Fourier space:

$$\delta(\boldsymbol{k},\tau) = \int \mathrm{d}^3q\,\mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{q}}\left[\mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{\Psi}(\boldsymbol{q},\tau)} - 1\right] = \int \mathrm{d}^3q\,\mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{q}} \sum_{m=1}^{\infty} \frac{\left[-i\boldsymbol{k}\cdot\boldsymbol{\Psi}(\boldsymbol{q},\tau)\right]^m}{m!}\,, \tag{2.36}$$

where we have used Eq. (2.35) to transform the integration variable from $\boldsymbol{x}$ to $\boldsymbol{q}$. The right-hand side clearly shows that even a linear displacement $\boldsymbol{\Psi}^{(1)}$ contributes at all orders to δ. In particular, at linear order (and expressed in configuration space again) we get

$$\nabla_q \cdot \boldsymbol{\Psi}^{(1)}(\boldsymbol{q},\tau) = -D(\tau)\,\delta_0(\boldsymbol{q})\,, \tag{2.37}$$

which is known as the Zel'dovich approximation (Zel'dovich, 2009), and can be interpreted as setting particles on straight trajectories with their velocities determined by the primordial density fluctuations. Higher-order corrections to the displacement field are obtained by solving an analogue to the Euler equation with the expansion

$$\boldsymbol{\Psi}(\boldsymbol{q},\tau) = \boldsymbol{\Psi}^{(1)}(\boldsymbol{q},\tau) + \boldsymbol{\Psi}^{(2)}(\boldsymbol{q},\tau) + \boldsymbol{\Psi}^{(3)}(\boldsymbol{q},\tau) + \ldots\,, \tag{2.38}$$

[5] At this time, τ_i, we must have $\boldsymbol{\Psi}(\boldsymbol{q},\tau_i) = 0$, which can be considered as a zeroth order approximation.

where each consecutive order contains an additional power of the linear displacement field, i.e. $\boldsymbol{\Psi}^{(2)} \sim \left(\boldsymbol{\Psi}^{(1)}\right)^2$, etc. Assuming, as before, that there are no velocity curl modes, it can further be shown that the displacement field is potential up to second order (Buchert et al., 1994; Catelan, 1995), in which case we write $\boldsymbol{\Psi}^{(n)}(\boldsymbol{q},\tau) = D_n(\tau)\,\nabla_q \phi^{(n)}(\boldsymbol{q})$ with $D_n(\tau)$ being the n-th order growth factor[6]. The solutions for the two potentials are (Buchert et al., 1994; Catelan, 1995):

$$\nabla_q^2 \phi^{(1)}(\boldsymbol{q}) = -\delta_0(\boldsymbol{q})\,, \tag{2.39}$$

$$\nabla_q^2 \phi^{(2)}(\boldsymbol{q}) = -\left[\left(\nabla_{q,ij}\phi^{(1)}(\boldsymbol{q})\right)\left(\nabla_{q,ji}\phi^{(1)}(\boldsymbol{q})\right) - \left(\nabla_q^2\phi^{(1)}(\boldsymbol{q})\right)^2\right] \equiv -\mathscr{G}_2(\phi^{(1)}(\boldsymbol{q}))\,, \tag{2.40}$$

where we have defined the second-order Galilean operator $\mathscr{G}_2$ that will make frequent appearances in Chap. 7. Similar to what we observed for the second-order Eulerian perturbation theory kernel, this operator encodes the effects from gravitational tides.

Finally, we can write the particle positions and velocities as (with $f_n \equiv \mathrm{d}\ln D_n/\mathrm{d}\ln a$):

$$\boldsymbol{x}(\boldsymbol{q},\tau) = \boldsymbol{q} + D_1(\tau)\,\nabla_q\phi^{(1)}(\boldsymbol{q}) + D_2(\tau)\,\nabla_q\phi^{(2)}(\boldsymbol{q})\,, \tag{2.41}$$

$$\boldsymbol{v}(\boldsymbol{q},\tau) = \frac{\mathrm{d}}{\mathrm{d}\tau}\boldsymbol{x}(\boldsymbol{q},\tau) = \mathscr{H}(\tau)\,f_1\,D_1(\tau)\,\nabla_q\phi^{(1)}(\boldsymbol{q}) + \mathscr{H}(\tau)\,f_2\,D_2(\tau)\,\nabla_q\phi^{(2)}(\boldsymbol{q})\,, \tag{2.42}$$

which forms the basis for setting up the initial conditions in N-body simulations. It was demonstrated in Crocce and Scoccimarro (2006a) that the second-order correction is especially important in order to suppress transients that get introduced by the Zel'dovich approximation, whereas third and higher-order only bring about negligible improvements.

2.3 Towards the smallest scales: halos and galaxies

Higher-order perturbation theory enables us to follow the evolution of initial density perturbations into the quasi-linear regime. Eventually, the density fluctuations will have grown so much that $\delta \gtrsim 1$, after which any perturbative treatment inevitably breaks down, and this usually happens long before galaxies start to form. At this point we could turn to N-body simulations that solve the exact particle dynamics under the influence of their mutual gravitational forces. However, in order to simulate the formation of galaxies, we would also have to account for a range of complicated (non-gravitational) processes that are still not understood in detail (see Mo et al., 2010, for an overview). Based on only a few physical assumptions, it is fortunately possible to gain an understanding of the abundance and distribution of dark matter halos that host the galaxies, as well as their relation. Alternatively, one can adopt an agnostic point of view about the small-

[6] By definition, $D_1(\tau) \equiv D(\tau)$, whereas to good accuracy for ΛCDM models $D_2(\tau) \approx -3/14\,D(\tau)^2$ (Bouchet et al., 1995).

scale physics and model the relation between galaxies and the large-scale matter field perturbatively, which requires the introduction of a range of unknown bias parameters, however.

2.3.1 Spherical Collapse

To gain some insight into the behaviour of density perturbations in the deeply nonlinear regime, let us consider a single, spherical over-density Δ embedded in an otherwise flat and matter-dominated universe. The Birkhoff theorem guarantees that its evolution is independent from the surrounding spacetime and only determined by the mass enclosed within. It can therefore be considered as an entirely separate universe, albeit with positive spatial curvature, and the corresponding Friedmann equations have a simple parametric solution for the radial size R of the perturbation at time t:

$$R(\vartheta) = \frac{R_{\mathrm{ta}}}{2}(1-\cos\vartheta)\,, \quad t(\vartheta) = \frac{t_{\mathrm{ta}}}{\pi}(\vartheta-\sin\vartheta)\,, \quad \text{with} \quad \vartheta \in \left[0, 2\pi\right]. \tag{2.43}$$

We see that as time goes on, the over-density first begins to expand, but eventually reaches its maximal extent at $\vartheta = \pi$ — the turn-around radius R_{ta}, which depends on the mass and the initial size of the perturbation — and subsequently collapses to a singularity of vanishing size. As the mass within the over-density must be conserved, the over-density itself is changing and based on Eq. (2.43) one can show that

$$\Delta(t) = 1 + \frac{9}{2}\frac{(\vartheta-\sin\vartheta)^2}{(1-\cos\vartheta)^3}\,. \tag{2.44}$$

In the linear regime we know from Sec. 2.2.2 that all perturbations are growing proportionally to the growth factor of the background, which in the present Einstein de-Sitter case scales as $D(t) = a(t) \propto (t/t_{\mathrm{ta}})^{2/3}$. For $\vartheta \ll 1$ the linear over-density must match with Eq. (2.44), which fixes the proportionality constant, so that

$$\Delta_0(t) = \frac{3}{20}\left(6\pi\,\frac{t}{t_{\mathrm{ta}}}\right)^{2/3} = \frac{3}{20}\,(6\vartheta - 6\sin\vartheta)^{2/3}\,. \tag{2.45}$$

If we now expand the fully nonlinear solution up to second order in terms of Δ_0 we find

$$\Delta(t) = \Delta_0(t) + \frac{17}{21}\,\Delta_0(t)^2 + \ldots\,, \tag{2.46}$$

which reproduces the coefficient we attributed to spherical collapse in the second-order Eulerian perturbation theory kernel (Eq. 2.33).

At turn-around time the linear density perturbation becomes roughly unity, so that any perturbative expansion is going to break down for $\vartheta \gtrsim \pi$. However, the subsequent collapse into a singularity is not physical either, as in reality slight departures from

spherical symmetry and other dissipative effects will cause the kinetic energy building up during the collapse to be converted into velocity dispersion, that is, random motions of individual particles. The system will then virialise and form a bound structure that is called a "halo". According to the virial theorem this process is complete once the time average of the potential energy equals twice that of the kinetic energy. That is the case when $R_{\rm vir} = R_{\rm ta}/2$ and assuming this occurs at $\vartheta = 2\pi$ one can work out that $\Delta_{\rm vir} \simeq 178$. In comparison, the linearly extrapolated density perturbation from Eq. (2.45) at this time is

$$\delta_{\rm cr} \equiv \Delta_0(t_{\rm vir}) \simeq 1.686\,, \tag{2.47}$$

which we define as the critical collapse density.

2.3.2 Halo mass function and halo bias

From the spherical collapse model we have obtained a criterion for forming a halo, namely whenever the linearly extrapolated over-density exceeds $\delta_{\rm cr}$. Can we use this approximation to predict the abundance of dark matter halos at any given time?

In order to identify halos of mass M or above in the field of linear density fluctuations δ_0, extrapolated to some time of interest, we filter it on a scale R (for instance, by applying a spherical top-hat window), which is related to the mass via the background density, i.e. $M = (4\pi/3)\,\overline{\rho}\,R^3$. We assume that all points at which the value of this smoothed field exceeds the critical collapse density are contained in a halo. Therefore, the total fractional volume enclosed in halos of size R or bigger is given by integrating the probability density function (PDF) of the smoothed density fluctuations $\delta_{0,R}$ over all values larger than $\delta_{\rm cr}$:

$$F_V(\delta_{0,R} > \delta_{\rm cr}) = \frac{1}{\sqrt{2\pi\,\sigma^2(R)}} \int_{\delta_{\rm cr}}^{\infty} {\rm d}\delta\, \exp\left[-\frac{\delta^2}{2\sigma^2(R)}\right]. \tag{2.48}$$

In accordance with the generic inflation prediction from Sec. 2.1.2 we have assumed here that $\delta_{0,R}$ follows Gaussian statistics, which are completely characterised by the variance on scale R,

$$\sigma^2(R) \equiv \langle \delta_{0,R}^2 \rangle = \int_{\boldsymbol{k}} P_0(k)\, W_R^2(k)\,, \tag{2.49}$$

where $W_R(k)$ is the smoothing window function in Fourier space. It was suggested by Press and Schechter (1974) that the volume fraction above should be the same as the total mass fraction $F_M(>M)$ contained in halos with mass greater than M. As the latter is determined by the halo mass function $\overline{n}_h(M)$ — the comoving number density of

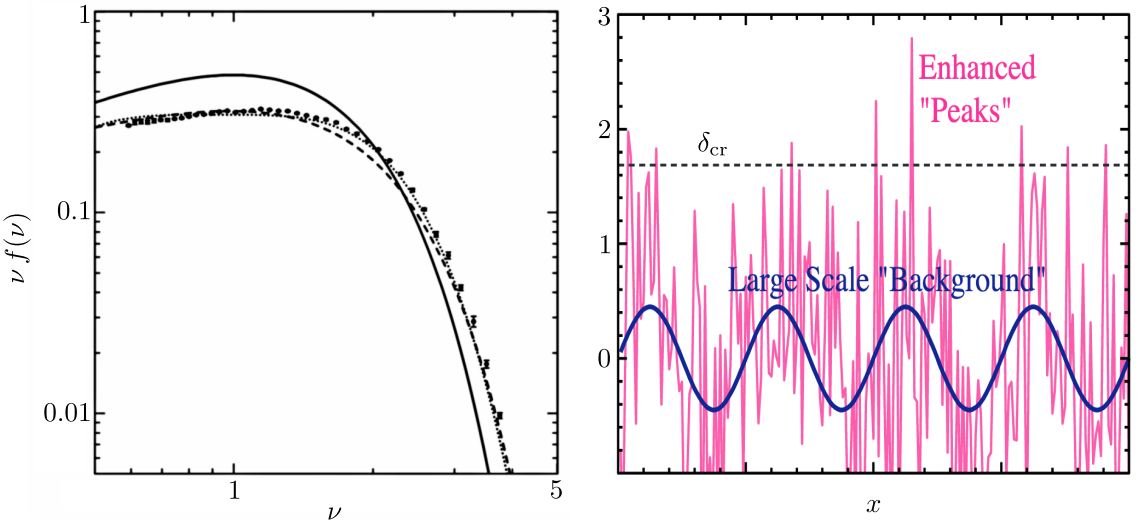


Fig. 2.3 *Left:* Multiplicity function $\nu\, f(\nu)$ as a function of the peak significance (from Zentner, 2007). The solid line is the Press-Schechter prediction from Eq. (2.51), whereas the dashed and dotted lines correspond to the Sheth-Tormen (Sheth and Tormen, 1999) and Jenkins (Jenkins et al., 2001) models. The data points are measurements from an N-body simulation by J. L. Tinker. *Right:* Illustration of the peak-background split idea (from Hu).

halos per logarithmic mass bin — we must have

$$F_M(>M) = \frac{1}{\overline{\rho}} \int_M^\infty \mathrm{d}\ln M'\, M'\, \overline{n}_h(M') = F_V(\delta_{0,R} > \delta_{\rm cr})\,. \tag{2.50}$$

Differentiating Eq. (2.50) with respect to mass yields the so-called Press-Schechter mass function (Press and Schechter, 1974)

$$\overline{n}_h(M) \equiv \frac{\mathrm{d}F_M(>M)}{\mathrm{d}M} = \frac{\overline{\rho}}{M}\,\nu\, f(\nu) \left|\frac{\mathrm{d}\ln\nu}{\mathrm{d}\ln M}\right|, \quad \text{with} \quad \nu\, f(\nu) = \sqrt{\frac{2}{\pi}}\,\nu\, \mathrm{e}^{-\nu^2/2}, \tag{2.51}$$

expressed in terms of the peak significance $\nu \equiv \delta_{\rm cr}/\sigma(R)$, which is large for very massive halos, and $\nu \ll 1$ for low-mass halos. From Eq. (2.51) we see that this implies that the number density of small halos displays a power-law behaviour, whereas halos more massive than $M_\star \simeq 2\times 10^{13}\, h^{-1}\, M_\odot$ (equivalent to $\nu = 1$) become exponentially rare. Moreover, if the halo mass function can be expressed in terms of the multiplicity function $\nu\, f(\nu)$, like in Eq. (2.51), its functional form is called universal, since all dependence on cosmological parameters, as well as on the redshift is contained in ν. A plot of the multiplicity function in comparison to measurements from an N-body simulation is shown in the left panel of Fig. 2.3.

The Press-Schechter approach reproduces the qualitative features of halo mass functions measured from N-body simulations, but leaves certain issues unaddressed. For instance, it does not account for the cloud-in-cloud problem (Peacock and Heavens, 1990), that is, the possibility that over-densities smaller than $\delta_{\rm cr}$ at some scale R can

still be part of a larger halo with $R' > R$. These over-densities are not accounted for in Eq. (2.50), which means that the mass contained in collapsed objects is underestimated[7]. A solution to this problem is provided by the excursion set formalism (Bond et al., 1991), which defines that a point can only belong to a halo of size R, if R is the *maximum* smoothing scale at which $\delta_{0,R}$ exceeds $\delta_{\rm cr}$. Another extension of the Press-Schechter approach is to consider mass-dependent collapse densities, as they are predicted if the assumption of spherical symmetry is relaxed (Sheth and Tormen, 2002; Sheth et al., 2001). The well-known Sheth-Tormen mass function is a variant of such an extension (Sheth and Tormen, 1999) and provides a much better fit to the results from N-body simulations than Press-Schechter (see also Fig. 2.3).

Although the halo mass function represents the mean number density of halos, we can also determine their local abundance by evaluating how $\overline{n}_h$ responds to a change in the background matter density. For that purpose let us again consider some constant spherical perturbation Δ, which can either correspond to an over-dense or under-dense region, and can be regarded as a separate universe with altered background density $\overline{\rho}^* = (1+\Delta)\,\overline{\rho}$. The response of the halo mass function is defined as the peak-background split bias parameter (Bardeen et al., 1986; Kaiser, 1984; Schmidt et al., 2013)

$$b \equiv \frac{1}{\overline{n}_h|_{\Delta=0}} \left.\frac{\partial \overline{n}_h|_\Delta}{\partial \Delta}\right|_{\Delta=0}\,, \tag{2.52}$$

where $\overline{n}_h|_{\Delta=0}$ denotes the halo mass function in the original universe. If we further promote Δ to a spatially varying field (for instance the linear matter fluctuations themselves) we can consider $\overline{n}_h|_\Delta$ as the local abundance of halos, $n_h(\boldsymbol{x})$. Given that $\Delta(\boldsymbol{x}) \ll 1$, Eq. (2.52) thus predicts a relation between the matter and halo fluctuation fields:

$$\delta_h(\boldsymbol{x}) \equiv \frac{n_h(\boldsymbol{x})}{\overline{n}_h} - 1 = b\,\Delta(\boldsymbol{x})\,, \tag{2.53}$$

which explains the terminology — halos are a *biased* tracer of the underlying (large-scale) matter field. For the spherical collapse model a uniform change in the background density implies a shift in the collapse threshold (Mo and White, 1996)

$$\delta_{\rm cr} \longrightarrow \delta_{\rm cr} - \Delta\,, \tag{2.54}$$

so that we should get a higher number of halos if Δ corresponds to an over-dense region and vice versa. This is portrayed by the right-hand plot of Fig. 2.3, which decomposes the density fluctuations into a large-scale background mode and its small-scale (non-linear) features. One can easily see that the collapse criterion is predominantly exceeded in regions where the background is large. When applying Eq. (2.52) to a universal mass

[7] This required the introduction of a "fudge"' factor of 2 in Eq. (2.51) (Press and Schechter, 1974).

function of the form given in Eq. (2.51) it follows that

$$b(M) = -\frac{1}{\sigma(M)} \frac{\mathrm{d}\ln\left[\nu f(\nu)\right]}{\mathrm{d}\nu}, \tag{2.55}$$

which has been evaluated for various mass functions calibrated against simulations and Hoffmann et al. (2015); Manera et al. (2010) find the agreement with the measured bias parameters to be at the order of $10-20\,\%$. However, as argued by Schmidt et al. (2013) these inaccuracies are not due to Eq. (2.52), but rather because of the calibration of the halo mass function, as well as the limited validity of Eq. (2.54), which is only an approximation in the case of aspherical halo formation. On the other hand, Eq. (2.52) can be implemented exactly, by running separate universe simulations, where the background quantities have been changed accordingly (Lazeyras et al., 2016; Wagner et al., 2015).

2.3.3 Galaxy bias from halo occupation distributions

If dark matter halos are biased tracers of the underlying matter fluctuations, we can expect the same to be true for galaxies, which reside in these halos. However, as galaxy formation involves the hydrodynamics of gas flows, feedback effects from star formation and supernovae etc., the spherical, and in particular pressure-less, collapse model from Sec. 2.3.1 cannot be applicable in this case. An attempt of describing the connection between dark matter halos and galaxies without the need to run complicated simulations, are the halo occupation distribution (HOD) models that are based on only a few physical insights.

The main idea of the HOD framework (Berlind and Weinberg, 2002; Kravtsov et al., 2004) is that each halo can host central and satellite galaxies, whose number depends in the simplest models solely on the mass M of the halo. The probability of hosting a central galaxy usually follows a softened step function that cuts off below M_{min}, while the satellite fraction is parametrised by a power law. In the model introduced by Zheng et al. (2007) the average number of centrals and satellites are therefore given by

$$\langle N_c(M)\rangle = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\ln M - \ln M_{\mathrm{min}}}{\sigma_{\ln M}}\right)\right], \tag{2.56}$$

$$\langle N_s(M)\rangle = \langle N_c(M)\rangle \times \begin{cases} \left(\frac{M-M_0}{M_1}\right)^{\alpha}, & M > M_0 \\ 0, & M \leq M_0 \end{cases}, \tag{2.57}$$

where M_0 is the minimum mass for a halo in order to host at least one satellite galaxy, $M_1 + M_0$ is the typical mass with exactly one such satellite, and $\sigma_{\ln M}$ represents a logarithmic scatter between galaxy luminosity and halo mass. The HOD is implemented on top of N-body simulations whose dark matter halos get populated with central and

satellite galaxies, drawn from a nearest-integer and Poisson distribution with the above means, respectively. In addition, satellite galaxies can only exist in halos which host a central galaxy, and their spatial distribution is assumed to follow the Navarro-Frenk-White density profile (Navarro et al., 1996) of the halo[8].

Based on these prescriptions we can compute the mean comoving number density of galaxies as an integral over the halo mass function:

$$\overline{n}_g = \int \mathrm{d}\ln M\, \overline{n}_h(M) \left[\langle N_c(M)\rangle + \langle N_s(M)\rangle\right], \tag{2.58}$$

where the parameter $M_{\min}$ can be adjusted so that Eq. (2.58) matches the number density of some observed population of galaxies. The remaining HOD parameters are fixed in such a way that the model reproduces their clustering properties (see Sec. 3.1), which works quite well over a large range of scales (Guo et al., 2015; Zehavi et al., 2011; Zu and Mandelbaum, 2015). Once the parameters are fixed the HOD model makes predictions about the large-scale galaxy bias. Applying the peak-background split argument (Eq. 2.52) to galaxies and assuming that the occupation statistics are invariant under a change in the background density, we get

$$b_g = \frac{1}{\overline{n}_g} \int \mathrm{d}\ln M\, \overline{n}_h(M) \left[\langle N_c(M)\rangle + \langle N_s(M)\rangle\right] b(M), \tag{2.59}$$

showing that the galaxy bias is inherited from their host dark matter halos. As the HOD parameters depend on the properties of the particular galaxy sample, such as their luminosities, so does the galaxy bias. For instance, more luminous galaxies tend to be found in more massive halos (Guo et al., 2015; Zehavi et al., 2011), which are thus more heavily biased according to Eq. (2.59).

2.3.4 Perturbative approach to galaxy bias

While phenomenological HOD models offer the possibility of describing galaxy fluctuations in the deeply nonlinear regime, they do rely on a number of simplifying parametrisations. These include the usual assumption that the occupation statistics only depend on the mass of the host halo, as well as that the dark matter halo fluctuations themselves are unaltered by the presence of baryonic matter. An alternative approach to modelling the relation between the galaxy and underlying matter fields, at least on large scales, aims to use a perturbative expansion and is thus more in spirit with the concepts from Sec. 3.2.1.

Based on the peak-background split idea (Bardeen et al., 1986; Kaiser, 1984) we were able to show that the galaxy perturbations are linearly related to those of the matter field, $\delta_g(\boldsymbol{x}) = b_g\,\delta(\boldsymbol{x})$. Yet, this can only be true on the largest scales because we had

[8] In practice they are assigned the positions of random dark matter particles within a halo.

to require that $\delta(\boldsymbol{x}) \ll 1$, so how do we proceed if we intend to adopt this approach in the quasi-linear regime? A natural extension, as suggested by Coles (1993); Gaztanaga and Frieman (1994), would be to expand the galaxy over-density in a Taylor-like series, which includes increasing powers of the matter fluctuations,

$$\delta_g(\boldsymbol{x}) = \sum_{n=1} \frac{b_n}{n!} \delta(\boldsymbol{x})^n \,, \tag{2.60}$$

and defines a whole set of new bias parameters b_n, with $b_1 \equiv b_g$. In light of our previous discussion in Sec. 2.3.1 we realise that this series expansion incorporates the effect from spherical collapse (which can be purely expressed in terms of powers of δ_0) on the galaxy fluctuations at each order of perturbation theory. Spherical collapse is however not the only process that determines the evolution of matter perturbations, so we can generalise Eq. (2.60) by allowing for all possible effects from the large-scale environment. We already identified two additional effects that appear at second order of perturbation theory in Sec. 2.2.3: a bulk flow due to the matter velocity field, and an impact from the tidal field. Including the former in the bias relation corresponds to the introduction of velocity bias, which would imply that galaxies are not comoving with the matter field. From simulations it was shown that the velocity fields of matter and halos differ only at the 2 % level (Zheng et al., 2015), so we can take this effect to be subdominant on the scales we are interested in here. The tidal field, on the other hand, is likely to leave a relevant imprint on the galaxy distribution, as argued in Baldauf et al. (2012); Catelan et al. (2000); Chan et al. (2012); McDonald and Roy (2009). For that reason we write the galaxy fluctuations up to second order as follows:

$$\delta_g(\boldsymbol{x}) = b_1 \delta(\boldsymbol{x}) + \frac{b_2}{2} \delta(\boldsymbol{x})^2 + \gamma_2 \mathcal{G}_2(\Phi|\boldsymbol{x}) + \dots \,, \tag{2.61}$$

where $\mathcal{G}_2$ (defined in Eq. 2.40) represents the tidal field with associated bias parameter γ_2. A convenient method for determining all gravity induced effects that would have to be included at even higher orders of perturbation theory will be presented in Chap. 7.

Each of the bias parameters is freely adjustable and must be marginalised over in an analysis of galaxy survey data. Therefore, they absorb all complexities of how the small-scale physics of galaxy formation responds to the large-scale environment. This, together with the fact that gravity alone determines the large-scale effects, guarantees the validity and robustness of the perturbative approach. It also highlights the fundamental difference to the HOD framework: where the perturbative treatment is agnostic about small-scale processes, the HOD, in contrast, relies on a concrete (and simplified) model. For the same reason, perturbative bias models can be readily applied to a whole range of different tracers of the matter field, such as quasars, the Lyman-α forest, and 21 cm emission. They can even accommodate naturally for assembly bias — the possi-

bility that the bias of the galaxies' host halos depend on other properties besides halo mass. If the galaxies inherit this assembly bias from their host halos, as explained in Sec. 2.3.3, their properties might become correlated with the halo formation history. In the perturbative model the bias parameters then become "effective" parameters, averaged over the bias of all galaxy types that were selected according to particular properties, such as their colour, spin etc. (see Sec. 9.2 of Desjacques et al., 2018). On the other hand, assembly bias is more difficult to implement in the HOD approach, which assumes that the number of central galaxies and satellites depends only on halo mass, but extensions have for instance been studied in Hearin et al. (2016).

Chapter 3

Measuring and predicting the clustering of galaxies

The previous chapter demonstrated that the origin of the observed large-scale structure can be understood based on the gravitational amplification of primordial inhomogeneities. How can we test the validity of this paradigm given that the only available data are the positions of galaxies? Clearly, it is not practical to probe evolving structures by studying individual systems because of the immense cosmic timescales, so as became clear already at the time of Hubble's observations, what is required is a statistical approach. The goal of such a description is to quantify the clustering of galaxies, which does not rely on individual objects and can be measured both as a function of time (via the galaxy redshifts) and spatial scale. For that reason this approach should provide valuable information on the process of structure formation as well as the background cosmological model.

The most notable measure for clustering is the two-point galaxy correlation function. It was first applied to the data from early galaxy catalogues by Katz and Mulders (1942); Limber (1954); Neyman and Scott (1952); Zwicky (1942) (amongst others), and remains the standard analysis tool in large-scale structure cosmology until today. That is because its measurement is relatively straight-forward, but more importantly, it can be robustly predicted from perturbation theory. Moreover, its Fourier transform is directly related to the spectrum of linear fluctuations, which implies it should contain all information about the large-scale distribution if the linear perturbations are indeed Gaussian as predicted by single-field inflation (which we will assume to be the case throughout). However, non-linear evolution and galaxy bias induce higher-order correlations, which will be of particular focus in this chapter.

We will begin by formally defining correlation functions of the galaxy distribution and derive how they are related to their Fourier space analogues, the N-point spectra. We compute theoretical predictions based on non-linear perturbation theory and a bias model developed in Sec. 2.2.3 and 2.3.4, which are useful to demonstrate how

three-point statistics break degeneracies between various cosmological parameters. Finally, we consider the gain in signal-to-noise from three-point correlations at the expense of a more complex covariance matrix, and finish with prospects of forthcoming surveys.

3.1 Statistical description of density fluctuations

According to the theory of inflation the inhomogeneities in the matter density are due to quantum fluctuations in the inflaton field and thus the outcome of a random process. We can therefore imagine our Universe to be a realisation of some statistical ensemble, such that all information about the density perturbations $\delta(\boldsymbol{x},\tau)$ are encoded in the probability density function (PDF) $\mathscr{P}$. If we take δ to be a continuous field (which can either refer to matter, halos, or galaxies) the PDF expresses the probability,

$$\mathscr{P}\big[\delta(\boldsymbol{x}_1),\ldots,\delta(\boldsymbol{x}_n)\big]\,\mathrm{d}\delta(\boldsymbol{x}_1)\cdots\mathrm{d}\delta(\boldsymbol{x}_n)\,, \tag{3.1}$$

for the field to have values between $\delta(\boldsymbol{x}_i)$ and $\delta(\boldsymbol{x}_i)+\mathrm{d}\delta(\boldsymbol{x}_i)$ with $i=1,\ldots,n$. Observational evidence indicates that our Universe appears homogeneous and isotropic after averaging over suitably large length scales. In the absence of a dynamical explanation, these observational results were elevated to the cosmological principle by Edward Milne (Milne, 1933, 1934), implying that these properties of our Universe are not to be thought of as exceptional, but rather shared amongst the entire statistical ensemble. This allows us to impose conditions on the PDF since the probability of having a certain field configuration should then be invariant under arbitrary translations and rotations:

$$\mathscr{P}\big[\delta(\boldsymbol{x}_1+\boldsymbol{\Delta}),\ldots,\delta(\boldsymbol{x}_n+\boldsymbol{\Delta})\big]=\mathscr{P}\big[\delta(\boldsymbol{x}_1),\ldots,\delta(\boldsymbol{x}_n)\big]\,, \tag{3.2}$$

$$\mathscr{P}\big[\delta\left(\mathsf{R}\,\boldsymbol{x}_1\right),\ldots,\delta\left(\mathsf{R}\,\boldsymbol{x}_n\right)\big]=\mathscr{P}\big[\delta(\boldsymbol{x}_1),\ldots,\delta(\boldsymbol{x}_n)\big]\,, \tag{3.3}$$

where $\boldsymbol{\Delta}$ denotes a shift in position and R a rotation matrix. Equipped with these properties we can now proceed to analyse the PDF with regard to its cosmological information, which can be most conveniently extracted by taking various moments.

3.1.1 Two-point correlation function

As the density perturbations vanish on average, the simplest moment we can compute is the product of δ at two different positions, which defines the two-point correlation function

$$\xi(\boldsymbol{x}_1,\boldsymbol{x}_2)\equiv\langle\delta(\boldsymbol{x}_1)\,\delta(\boldsymbol{x}_2)\rangle\equiv\int\mathrm{d}\delta(\boldsymbol{x}_1)\,\mathrm{d}\delta(\boldsymbol{x}_2)\,\mathscr{P}\big[\delta(\boldsymbol{x}_1),\delta(\boldsymbol{x}_2)\big]\,\delta(\boldsymbol{x}_1)\,\delta(\boldsymbol{x}_2)\,. \tag{3.4}$$

Using statistical homogeneity, i.e. Eq. (3.2), we can show that the correlation function can only depend on the separation $\boldsymbol{r} = \boldsymbol{x}_1 - \boldsymbol{x}_2$ between the two positions, while statistical isotropy (Eq. 3.3) forbids any preferred directions, so that ξ must be completely characterised by the magnitude $r = |\boldsymbol{r}|$ alone.

To gain an understanding of how the two-point function relates to the clustering of point-like objects — for instance, dark matter particles or observed galaxies — we need to consider a *discrete* sampling of the continuous field. This is usually taken to be a Poisson point process, where the probability of finding an object in a volume element δV centred at position $\boldsymbol{x}$ is proportional to the density at that point. More precisely, the number of objects in δV is Poisson distributed with intensity $\rho(\boldsymbol{x})\,\delta V/m$, where m denotes the mass of the discrete tracer, i.e. of the dark matter particle or a typical galaxy. To distinguish from the ensemble average above we indicate an average over the Poisson distribution as $\langle\ldots\rangle_\bullet$, so that the mean number of objects per volume element at position $\boldsymbol{x}$ is given by

$$\langle n(\boldsymbol{x})\rangle_\bullet = \frac{\rho(\boldsymbol{x})}{m} = \overline{n}\left[1+\delta(\boldsymbol{x})\right], \tag{3.5}$$

and $\overline{n} = \langle\langle n(\boldsymbol{x})\rangle_\bullet\rangle$ stands for the ensemble averaged number density. The defining characteristic of a Poisson process is further that the probability of finding an object in one volume element is independent from that in another. The average number of objects at positions $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ is therefore the product of Eq. (3.5):

$$\begin{aligned}\langle n(\boldsymbol{x}_1)\,n(\boldsymbol{x}_2)\rangle_\bullet &= \begin{cases} \overline{n}^2\left[1+\delta(\boldsymbol{x}_1)\right]\left[1+\delta(\boldsymbol{x}_2)\right], & \boldsymbol{x}_1 \neq \boldsymbol{x}_2 \\ \overline{n}^2\left[1+\delta(\boldsymbol{x}_1)\right]^2 + \frac{\overline{n}}{\delta V}\left[1+\delta(\boldsymbol{x}_1)\right], & \boldsymbol{x}_1 = \boldsymbol{x}_2 \end{cases} \\ &= \overline{n}^2\left[1+\delta(\boldsymbol{x}_1)\right]\left[1+\delta(\boldsymbol{x}_2)\right] + \overline{n}\left[1+\delta(\boldsymbol{x}_1)\right]\delta_D(\boldsymbol{x}_1-\boldsymbol{x}_2)\,, \end{aligned} \tag{3.6}$$

where in the case that $\boldsymbol{x}_1 = \boldsymbol{x}_2$ we have used that the second moment of the Poisson distribution satisfies $\langle\mathcal{N}^2\rangle_\bullet = \langle\mathcal{N}\rangle_\bullet + \langle\mathcal{N}\rangle_\bullet^2$. Taking now the average over the ensemble and inserting the definition from Eq. (3.4) we get

$$\langle\langle n(\boldsymbol{x}_1)\,n(\boldsymbol{x}_2)\rangle_\bullet\rangle = \overline{n}^2\left[1+\xi(r)\right] + \overline{n}\,\delta_D(\boldsymbol{x}_1-\boldsymbol{x}_2)\,, \tag{3.7}$$

showing that the two-point correlation function describes the excess probability of finding pairs of galaxies with a given separation over a completely random distribution (i.e. a Poisson process with constant intensity). If $\xi > 0$ objects thus tend to be more clustered on average, whereas $\xi < 0$ indicates a deficiency of clustering. The second contribution in Eq. (3.7) is due to the counting of self-pairs, which is commonly referred to as shot noise. It can be easily ignored by restricting ourselves to separations $r > 0$,

but it will play a crucial role for the covariance of the two-point correlation function as we will see in Sec. 3.3.2.

If the density perturbations are Gaussian, their joint N-point PDF[1] is given by

$$\mathscr{P}(\delta_1,\ldots,\delta_n)\,\mathrm{d}\delta_1\cdots\mathrm{d}\delta_n = \frac{1}{\sqrt{(2\pi)^n\,\det\Xi}}\exp\left[-\frac{1}{2}\sum_{i,j=1}^{n}\delta_i\,\Xi_{ij}^{-1}\,\delta_j\right]\mathrm{d}\delta_1\cdots\mathrm{d}\delta_n\,, \tag{3.8}$$

where we have used the short-hand notation $\delta_i \equiv \delta(\boldsymbol{x}_i)$ and the matrix Ξ encodes the correlations between field values at various positions, that is, $\Xi_{ij} \equiv \xi(|\boldsymbol{x}_i - \boldsymbol{x}_j|)$. This highlights the importance of the two-point correlation function, as it completely determines the PDF of a Gaussian field and consequently contains all statistical information. Fluctuations probed via the CMB have not yet undergone non-linear evolution, so if the quantum fluctuations from inflation were Gaussian, it would therefore be sufficient to extract the two-point function from the CMB. Conversely, any detection of a higher-order correlation function (see next section) would immediately point towards the existence of primordial non-Gaussianity. However, we will see in Sec. 3.1.3 that this is no longer true after gravitational evolution of the density fluctuations as this leads to the appearance of an additional form of non-Gaussianity.

3.1.2 Higher-order correlation functions

For non-Gaussian fields the statistical information is spread out over the whole hierarchy of moments. An example that illustrates the importance of measuring statistics beyond the two-point correlation function in these situations is shown in Fig. 3.1. It displays the distribution of galaxies from a Sloan Digital Sky Survey (SDSS) mock catalogue on the left-hand side, confronted with a realisation of a Rayleigh-Lèvy flight on the right. A Rayleigh-Lèvy flight is a collection of individual random walks, all of which start from a random position. The direction of each step is also chosen at random, whereas its size is drawn from a distribution whose probability drops by a power law for step sizes beyond a fixed length threshold (see Peebles, 1980; Szapudi and Colombi, 1996, for details). Although the two distributions in Fig. 3.1 look clearly different, the parameters of the Rayleigh-Lèvy flight were adjusted such that both have the exact same two-point correlation function (Sefusatti and Scoccimarro, 2005). To disentangle different morphologies, for instance between compact and elongated structures, it is inevitable to turn to higher-order moments of the PDF.

These can be defined in exact analogy with Eq. (3.4), but what we will define as the higher-order correlation functions ξ_N are in fact the cumulants of $\mathscr{P}$, which have all

[1] Note that we have already made use of a Gaussian PDF with $n = 1$ in Sec. 2.3.2.

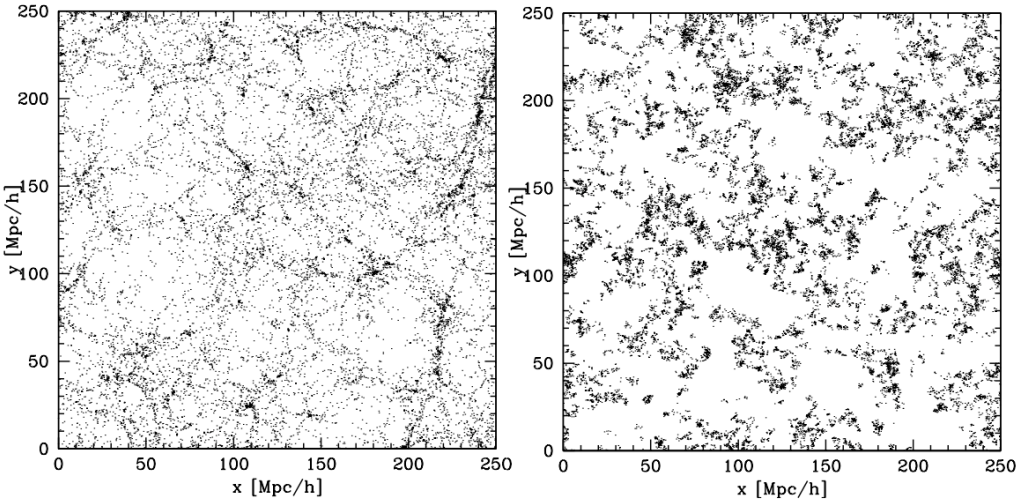


Fig. 3.1 Slices of thickness 50 Mpc/h through a mock galaxy distribution for SDSS (left panel) and a Rayleigh-Lèvy flight (right panel). The latter has a number of free parameters that were chosen such that both distribution have the same two-point correlation function, despite looking different by eye. (From Sefusatti and Scoccimarro, 2005).

contributions from products of lower-order cumulants subtracted off:

$$\begin{aligned}\xi_N(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N) &\equiv \langle\delta(\boldsymbol{x}_1)\cdots\delta(\boldsymbol{x}_N)\rangle_c \\ &\equiv \langle\delta(\boldsymbol{x}_1)\cdots\delta(\boldsymbol{x}_N)\rangle - \sum_{\mathscr{S}\in P\{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N\}}\prod_{s\in\mathscr{S}}\langle\delta(\boldsymbol{x}_{s_1})\cdots\delta(\boldsymbol{x}_{s_n})\rangle_c\,, \end{aligned} \tag{3.9}$$

where the sum runs over partitions of the set of positions $\{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N\}$, and s is a subset of length n contained in the partition $\mathscr{S}$. As we have that $\langle\delta(\boldsymbol{x})\rangle = 0$, all subsets must have length larger than one, which implies that the second and third moments are identical to their cumulants.

In this work we are particularly interested in the three-point correlation function $\zeta \equiv \xi_3$, which measures the probability of finding triplets of objects, in excess over a random distribution of points and their two-point correlations. Using the Poisson process introduced in Sec. 3.1.1 and distinguishing cases where either two of the positions or all three coincide, we get

$$\begin{aligned}\langle\langle n(\boldsymbol{x}_1)\,n(\boldsymbol{x}_2)\,n(\boldsymbol{x}_3)\rangle_\bullet\rangle &= \overline{n}^3\big[1+\xi(r_{12})+\xi(r_{23})+\xi(r_{31})+\zeta(r_{12},r_{23},r_{31})\big] \\ &\quad+\overline{n}^2\big[1+\xi(r_{12})\big]\,\delta_D(\boldsymbol{x}_2-\boldsymbol{x}_3)+\text{cyc.} \\ &\quad+\overline{n}\,\delta_D(\boldsymbol{x}_1-\boldsymbol{x}_2)\,\delta_D(\boldsymbol{x}_2-\boldsymbol{x}_3)\,. \end{aligned} \tag{3.10}$$

For notational convenience we have defined $r_{ij} \equiv |\boldsymbol{x}_i - \boldsymbol{x}_j|$ as the separation between two positions and statistical homogeneity and isotropy has allowed us to write the three-point correlator as a function of the three separations only. As for the two-point correlation function, we have obtained additional contributions from shot noise —

now due to the counting of self-pairs and self-triplets. If the continuous underlying field is Gaussian, $\zeta = 0$, and we see that the probabilities of finding triplets is entirely determined by ξ_2.

3.1.3 Clustering statistics in Fourier space

We have seen in Sec. 2.2.3 that a Fourier space description facilitates solving the fluid equations in the non-linear regime of structure formation. For that reason it is useful to record expressions for the correlation functions of Fourier modes $\delta(\boldsymbol{k})$, which we will later compute in the perturbation theory framework.

We have already defined the two-point correlator of Fourier modes as the power spectrum P in Sec. 2.1.2, but in addition we can now determine how it is related to the two-point correlation function ξ by plugging in their Fourier transformations:

$$\begin{aligned}\langle\delta(\boldsymbol{k}_1)\,\delta(\boldsymbol{k}_2)\rangle &= \int \mathrm{d}^3x_1\,\mathrm{d}^3x_2\,\mathrm{e}^{-i(\boldsymbol{k}_1\cdot\boldsymbol{x}_1+\boldsymbol{k}_2\cdot\boldsymbol{x}_2)}\,\langle\delta(\boldsymbol{x}_1)\,\delta(\boldsymbol{x}_2)\rangle \\ &= (2\pi)^3\,\delta_D(\boldsymbol{k}_{12})\int \mathrm{d}^3x_1\,\mathrm{e}^{-i\boldsymbol{k}_1\cdot\boldsymbol{x}_1}\,\xi(x_1) \\ &\equiv (2\pi)^3\,P(k_1)\,\delta_D(\boldsymbol{k}_{12})\,, \end{aligned} \tag{3.11}$$

To perform one of the integrations we have exploited statistical homogeneity, which has lead to the appearance of the Dirac delta distribution and implies that two different wave modes must be uncorrelated. Moreover, we notice that the power spectrum is simply the Fourier transform of the two-point correlation function, a result widely known as the Wiener-Khinchin theorem. That suggests we can directly compute the power spectrum for discrete tracers from the analogous result for ξ, and using Eq. (3.7) expressed in terms of perturbations in the number density, $\delta n(\boldsymbol{x}) \equiv n(\boldsymbol{x})/\overline{n} - 1$, gives

$$\langle\langle\delta n(\boldsymbol{k}_1)\,\delta n(\boldsymbol{k}_2)\rangle_{\bullet}\rangle = (2\pi)^3\left[P(k_1) + \frac{1}{\overline{n}}\right]\delta_D(\boldsymbol{k}_{12})\,, \tag{3.12}$$

which shows that the discrete power spectrum is modulated by a constant shot noise contribution on all scales k_1 (opposed to ξ). The Wiener-Khinchin theorem also applies to higher-order correlation functions, for instance, the three-point function ζ is the Fourier transform of the bispectrum B, which is defined as follows:

$$\langle\delta(\boldsymbol{k}_1)\,\delta(\boldsymbol{k}_2)\,\delta(\boldsymbol{k}_3)\rangle \equiv (2\pi)^3\,B(k_1,k_2,k_3)\,\delta_D(\boldsymbol{k}_{123})\,. \tag{3.13}$$

Statistical homogeneity dictates that the three Fourier modes of the bispectrum form a closed triangle, while the absence of preferred directions requires that B depends only on the lengths of the triangle sides, or alternatively on two of the sides and the angle between them etc. The discrete bispectrum may be readily obtained from Eq. (3.10), so

that

$$\langle\langle \delta n(\boldsymbol{k}_1)\,\delta n(\boldsymbol{k}_2)\,\delta n(\boldsymbol{k}_3)\rangle_{\bullet}\rangle = (2\pi)^3 \left\{ B(k_1,k_2,k_3) + \frac{1}{\overline{n}}\Big[P(k_1)+P(k_2)+P(k_3)\Big] + \frac{1}{\overline{n}^2}\right\} \times \delta_D(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)\,. \tag{3.14}$$

Similar to the power spectrum, shot noise impacts the discrete bispectrum on all scales, which means that its contributions have to be subtracted if we aim to isolate the signal of the continuous field from a measurement.

We can provide an interesting perspective on clustering in Fourier space by taking into account that the density fluctuations are real, and that the Fourier modes must therefore satisfy the reality constraint $\delta(\boldsymbol{k}) = \delta^*(-\boldsymbol{k})$. Combining this with the condition imposed by statistical homogeneity it follows that the power spectrum $P(\boldsymbol{k}) \sim \langle|\delta(\boldsymbol{k})|^2\rangle$ is a function of the Fourier amplitudes alone and consequently blind to all information stored in the Fourier phases. In contrast, the bispectrum and any other higher-order statistic derived from the correlation function hierarchy, contain a mixture of information coming from both. However, if the phases are randomly distributed, i.e. their PDF is uniform on the interval $[0,2\pi]$, it is easy to see that all higher-order statistics must vanish. Considering for example the bispectrum, an average over the contribution from the phases $\theta(\boldsymbol{k})$ gives:

$$\langle \delta(\boldsymbol{k}_1)\,\delta(\boldsymbol{k}_2)\,\delta(\boldsymbol{k}_3)\rangle = \langle|\delta(\boldsymbol{k}_1)|\,|\delta(\boldsymbol{k}_2)|\,|\delta(\boldsymbol{k}_3)|\rangle \times \underbrace{\int \prod_{i=1}^{3} \frac{\mathrm{d}\theta(\boldsymbol{k}_i)}{2\pi}\, \mathrm{e}^{i[\theta(\boldsymbol{k}_1)+\theta(\boldsymbol{k}_2)+\theta(\boldsymbol{k}_3)]}}_{=0}, \tag{3.15}$$

where we have used that the PDF of amplitudes and phases must be separable in this case. The same holds for any higher-order correlation of phase factors[2], so by means of the Edgeworth expansion we can conclude that a field with random phases has to be Gaussian. Conversely, detecting correlations in the phases is therefore a direct indicator for non-Gaussianity, which can be neatly visualised as is demonstrated in Fig. 3.2 (Coles and Chiang, 2000). The left-hand column shows the spatial distribution of particles in a two-dimensional N-body simulation, set up with Gaussian initial conditions, at an early and late time step. The next column displays the corresponding Fourier phases, represented by colour hue, while the last two show the difference of neighbouring phase factors, $\theta(k_x+1)-\theta(k_x)$, in x- and y-direction, respectively. At the later time step certain hues clearly dominate over others in the phase differences, which implies that the phases have become correlated through gravitational evolution. From this simple

[2] If there is an even number of phase factors, they can cancel each other in pairs given that $\boldsymbol{k}_1 = -\boldsymbol{k}_2$, and hence $\theta(\boldsymbol{k}_1) = -\theta(\boldsymbol{k}_2)$, etc. However, these terms correspond to products of two-point correlators, which have been subtracted from the higher-order correlation functions (see Eq. 3.9).

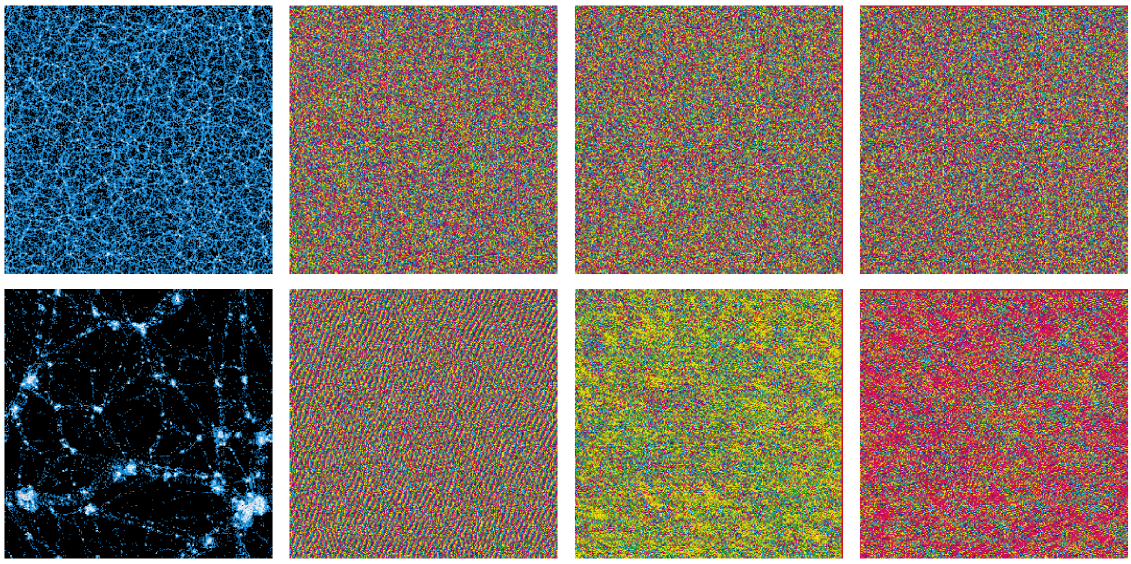


Fig. 3.2 The evolution of phase correlations in a two-dimensional N-body simulation with Gaussian initial conditions. First column: particle distribution at an initial (top) and late time step (bottom). Second column: Fourier phases of the particle distribution, visualised by colour hue. Third and fourth columns: differences between neighbouring phase factors in the x- and y-direction, respectively. (From Coles and Chiang, 2000).

observation we can already infer that the non-linear growth of structures introduces non-Gaussianity.

3.2 Clustering statistics from perturbation theory

The perturbation theory framework allows us to understand the emergence of non-Gaussianity from gravitational evolution. As we have seen in Sec. 2.2.3, non-linearities in the fluid equations lead to the coupling of modes, giving rise to corrections of the linear density perturbations $\delta^{(1)}(\boldsymbol{k},\tau) \equiv D(\tau)\,\delta_0(\boldsymbol{k})$ that scale as $\left(\delta^{(1)}\right)^n$ at n-th order in perturbation theory. Even if the linear fluctuations are Gaussian, they can correlate with these non-linear corrections and thus produce non-zero higher-order correlation functions.

3.2.1 Matter perturbations

To analyse the predictions from perturbation theory for the power spectrum and bispectrum let us begin by considering the matter fluctuations, before we incorporate the effects from galaxy bias in a separate step. The correlation of two linear modes defines the linear power spectrum

$$\langle \delta^{(1)}(\boldsymbol{k}_1,\tau)\,\delta^{(1)}(\boldsymbol{k}_2,\tau)\rangle = (2\pi)^3\,P_L(k_1,\tau)\,\delta_D(\boldsymbol{k}_{12})\,, \tag{3.16}$$

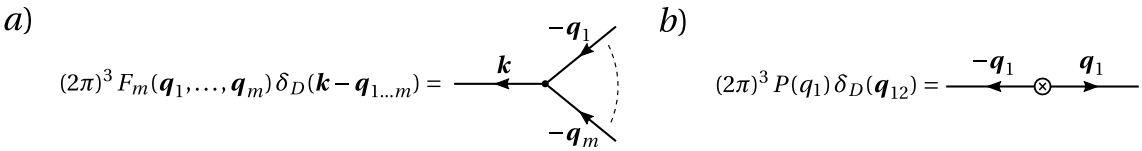
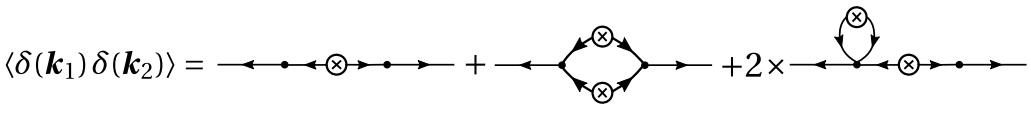


Fig. 3.3 Feynman rules for large-scale structure. The vertices (*a*) represent m-th order perturbation theory solutions $\delta^{(m)}(\boldsymbol{k})$, whereas crossed circles (*b*) indicate linear power spectra from pairing two internal lines.

where $P_L(k,\tau) \equiv D(\tau)^2\,P_0(k)$ is related to the spectrum of primordial fluctuations $P_0(k)$, given in Eq. 2.21. As we assume that $\delta^{(1)}$ is Gaussian, all higher-order correlation functions constructed from an odd number of linear density modes must vanish, which leaves the following series of terms upon plugging in the perturbative expansion (suppressing time arguments from now on):

$$\langle\delta(\boldsymbol{k}_1)\,\delta(\boldsymbol{k}_2)\rangle = \langle\delta^{(1)}(\boldsymbol{k}_1)\,\delta^{(1)}(\boldsymbol{k}_2)\rangle + \underbrace{\langle\delta^{(2)}(\boldsymbol{k}_1)\,\delta^{(2)}(\boldsymbol{k}_2)\rangle}_{\sim P_{22}(k_1)} + 2\,\underbrace{\langle\delta^{(3)}(\boldsymbol{k}_1)\,\delta^{(1)}(\boldsymbol{k}_2)\rangle}_{\sim P_{31}(k_1)} + \ldots \tag{3.17}$$

In analogy with Eq. (3.16) the two correction terms can be defined as power spectra P_{22} and P_{31}, both of which are fourth order in terms of $\delta^{(1)}$. To derive their explicit expressions we have to use the corresponding perturbation theory solutions from Eq. 2.30 and expand the ensemble average into products of linear power spectra by means of the Wick theorem. It is convenient to describe this procedure diagrammatically by introducing a set of "Feynman rules", whereby any N-point correlation function can be expressed as a series of diagrams with exactly n vertices. The vertices represent the m-th order solutions $\delta^{(m)}(\boldsymbol{k})$ and are therefore assigned the factor $(2\pi)^3\,F_m(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_m)\,\delta_D(\boldsymbol{k}-\boldsymbol{q}_{1\ldots m})$. They are depicted as the intersection of one outgoing line of momentum $\boldsymbol{k}$, and m internal lines with momenta $-\boldsymbol{q}_1,\ldots,-\boldsymbol{q}_m$, as shown in Fig. 3.3. Each internal line must further be connected to another one (either from the same or a different vertex), and any pair of lines, indicated by a crossed-out circle, contributes with $(2\pi)^3\,P_L(q_1)\,\delta_D(\boldsymbol{q}_{12})$ to the amplitude of the diagram. Finally, all internal momenta are integrated over and each diagram is multiplied by a symmetry factor that accounts for the possible number of connecting the internal lines. With these directions we can represent Eq. (3.17) as



(3.18)

and we see that the two non-linear corrections involve one internal loop each. Taken together they are categorised as the *one-loop contribution* to the matter power spectrum,

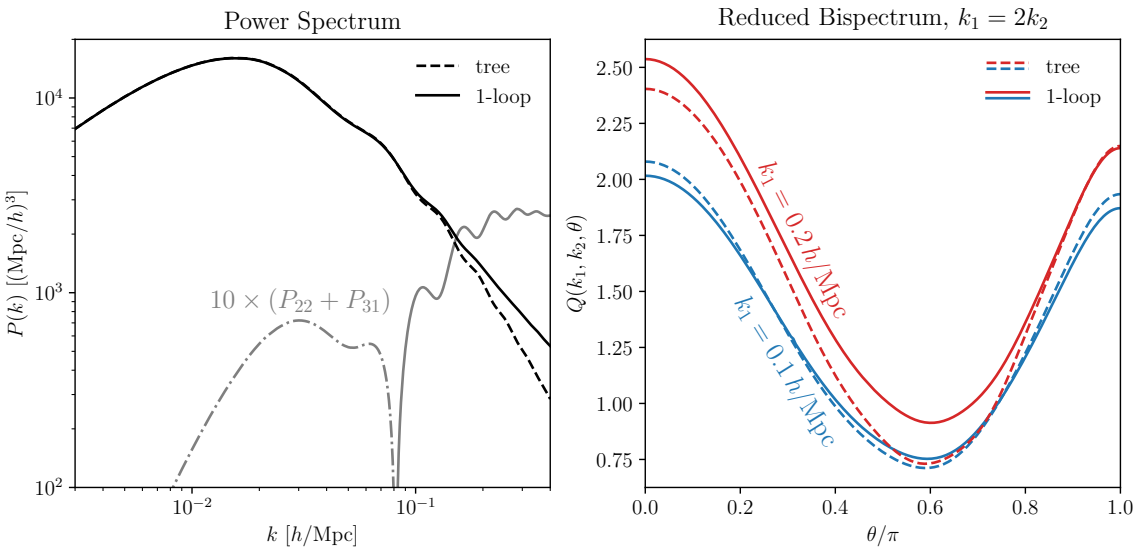


Fig. 3.4 Perturbation theory predictions for the power spectrum (left) and reduced bispectrum (right). Linear ("tree-level") results are shown as dashed lines, whereas the solid lines include one-loop contributions. The left panel also displays the one-loop corrections on their own, boosted by a factor of 10, whose negative sign at small k is indicated by the dot-dashed line. The reduced bispectrum is plotted as a function of the angle θ between $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ with the fixed constraint $k_1 = 2k_2$.

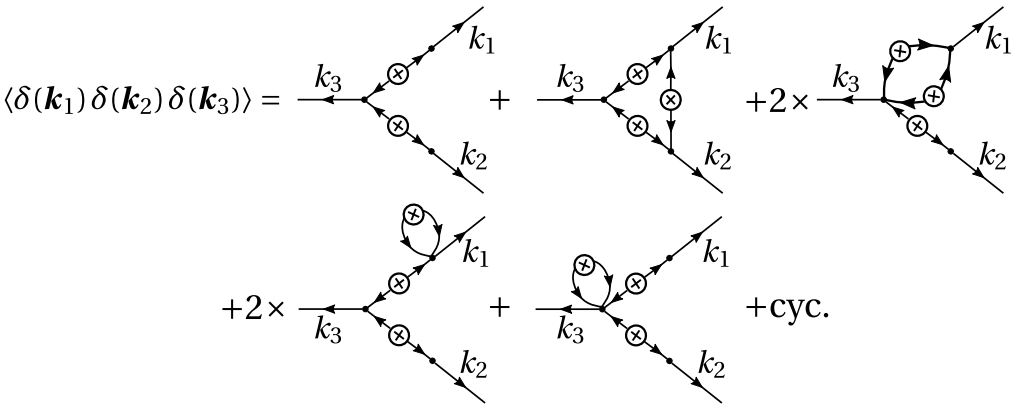
and following the rules above, the two diagrams are given by:

$$P_{22}(k) = 2\int_{\boldsymbol{q}} \left[F_2(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q})\right]^2 P_L(q)\, P_L(|\boldsymbol{k}-\boldsymbol{q}|)\,, \tag{3.19}$$

$$P_{31}(k) = 3\, P_L(k) \int_{\boldsymbol{q}} F_3(\boldsymbol{k}, \boldsymbol{q}, -\boldsymbol{q})\, P_L(q)\,. \tag{3.20}$$

The left-hand panel of Fig. 3.4 compares the linear power spectrum with and without the one-loop contributions for a ΛCDM model at redshift $z = 0.34$, showing that non-linear evolution leads to an enhancement of clustering on small scales (higher wave numbers k). These theory predictions have been well tested against numerical simulations (e.g. in Carlson et al., 2009; Jeong and Komatsu, 2006; Taruya et al., 2013), and setting an accuracy goal of ∼ 3 % these studies find that the one-loop power spectrum extends the regime of validity of linear theory only slightly from ∼ 0.1 h/Mpc to ∼ 0.12 h/Mpc at redshift zero. At higher redshifts non-linear perturbation theory generally works better, for instance at $z = 1$ it is accurate up to 0.15 h/Mpc. However, beyond these scales the one-loop power spectrum consistently over-estimates the clustering signal, which can be somewhat alleviated with more sophisticated modelling techniques, such as the renormalized perturbation theory (RPT) (Bernardeau et al., 2008; Crocce and Scoccimarro, 2006b, 2008; Taruya et al., 2012), or the effective field theory approach (EFT) (Baumann et al., 2012; Carrasco et al., 2012).

Applying the above rules to the bispectrum we can immediately write down its diagrammatic expansion up to the one-loop level,

$\langle \delta(\boldsymbol{k}_1)\,\delta(\boldsymbol{k}_2)\,\delta(\boldsymbol{k}_3)\rangle =$ 

(3.21)

where "cyc." stands for the two remaining permutations over k_1, k_2 and k_3. The lowest order, or tree-level, contribution comes from the correlation of two linear and a gravity induced second-order mode. The resulting bispectrum scales as P_L^2 and exhibits a strong configuration dependence (Fry, 1984; Fry et al., 1993) as it is directly proportional to the second-order perturbation theory kernel:

$$B_{211}(k_1,k_2,k_3) = 2F_2(\boldsymbol{k}_1,\boldsymbol{k}_2)\,P_L(k_1)\,P_L(k_2) + \text{cyc.} \tag{3.22}$$

From our discussion in Sec. 2.2.3 we know that the perturbation theory kernels are mostly cosmology invariant, so we can isolate the configuration dependence by computing the reduced bispectrum,

$$Q(k_1,k_2,k_3) \equiv \frac{B(k_1,k_2,k_3)}{P(k_1)\,P(k_2) + P(k_2)\,P(k_3) + P(k_3)\,P(k_1)}\,, \tag{3.23}$$

which is shown in the right-hand panel of Fig. 3.4. We plot $Q(k_1,k_2,k_3)$ as a function of the angle between $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ for triangles satisfying $k_1 = 2k_2$, and we note that collinear configurations ($\theta = 0,\pi$) display a particularly strong signal. This reflects the fact that gravitational instability generates large-scale bulk flows in preferred directions, namely along gradients of the density and velocity fields, which gives rise to more filamentary structures (Scoccimarro, 1997; Scoccimarro et al., 1998a). Mathematically, this is represented by the two non-linear terms in the continuity and Euler equations (see Eqs. 2.11 and 2.12), i.e. $\boldsymbol{v}\cdot\nabla\delta$ and $(\boldsymbol{v}\cdot\nabla)\boldsymbol{v}$, which together determine the scale dependence of F_2. On smaller scales bulk flows will start to interact with each other, making non-collinear clustering configurations more likely and thus flattening out the amplitude of Q. The one-loop bispectrum (Scoccimarro et al., 1998a), depicted by the solid lines, already gives an indication of this effect and thus improves on the tree-level prediction, albeit only for a limited range of scales. As analysed by Lazanu

et al. (2016), it tends to break down on scales comparable to those where the one-loop power spectrum begins to fail, whereas RPT and EFT methods offer the potential of significantly increasing its regime of validity (Baldauf et al., 2015; Bernardeau et al., 2012; Lazanu et al., 2016).

3.2.2 Galaxy perturbations

The Poisson probability of observing point-like objects is determined by fluctuations in the underlying density according to Eq. (3.5). When considering galaxies, this underlying field refers to the continuous galaxy over-density $\delta_g(\boldsymbol{x})$, which as we have seen in Sec. 2.3 is biased with respect to the matter perturbations. This bias has to be taken into account when comparing the measured power spectrum or bispectrum from a given galaxy population to theory predictions.

In the perturbative treatment of galaxy bias (Sec. 2.3.4) the galaxy fluctuations are proportional to those of the matter density on large scales, so that the linear galaxy power spectrum is simply given by

$$P_{g,L}(k) = b_1^2 \, P_L(k) \,. \tag{3.24}$$

Because the constant of proportionality — the linear bias parameter b_1 — is unknown a priori, it has to be determined from the data. Unfortunately, the bias factor is degenerate with the amplitude A_s of the linear matter power spectrum, which is often parametrised as the variance of fluctuations in spheres of radius 8 Mpc/h, linearly extrapolated to redshift zero:

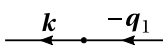
$$\sigma_8^2 \equiv \langle \left[\delta_W^{(1)}(\boldsymbol{x}, z=0)\right]^2 \rangle = \frac{1}{2\pi^2} \int_0^\infty \mathrm{d}k \, k^2 \, |W(k)|^2 \, P_L(k, z=0) \,, \tag{3.25}$$

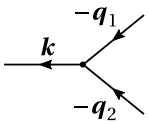
where $W(k)$ is the Fourier transform of the corresponding spherical top-hat smoothing function. This means a power spectrum measurement can only yield the combination $b_1 \, \sigma_8$.

Higher-order terms in the perturbative bias expansion scale as $(\delta^{(1)})^n$, like the non-linear corrections induced by gravitational evolution. Hence, these can couple to matter perturbations at all orders and therefore produce additional "bias corrections" either at loop- or even at tree-level. In the language of the Feynman rules above, we can consider the bias terms as vertices with their own kernel functions. Following Eq. 2.61

the possibilities at the first two orders extend to

1st) 2nd)

$$(2\pi)^3\, b_1\, \delta_D(\boldsymbol{k}-\boldsymbol{q}_1) =$$ 

$$\left.\begin{matrix} b_1\, F_2(\boldsymbol{q}_1,\boldsymbol{q}_2) \\ b_2/2 \\ \gamma_2\, K(\boldsymbol{q}_1,\boldsymbol{q}_2) \end{matrix}\right\} (2\pi)^3\, \delta_D(\boldsymbol{k}-\boldsymbol{q}_{12}) =$$ 

(3.26)

where $K(\boldsymbol{q}_1,\boldsymbol{q}_2) \equiv \left[\boldsymbol{q}_1\cdot\boldsymbol{q}_2/(q_1\,q_2)\right]^2 - 1$ denotes the tidal field operator $\mathscr{G}_2$ in Fourier space. While the power spectrum only receives corrections at the one-loop level (which we will analyse in detail in Chap. 7), the tree-level bispectrum involves a second-order vertex, which gives rise to two extra contributions that alter the configuration dependence:

$$\begin{aligned} B_{g,211}(k_1,k_2,k_3) = b_1^3\, B_{211}(k_1,k_2,k_3) + b_2\, b_1^2\left[P_L(k_1)\,P_L(k_2) + \text{cyc.}\right] \\ + 2\,\gamma_2\, b_1^2\left[K(\boldsymbol{k}_1,\boldsymbol{k}_2)\,P_L(k_1)\,P_L(k_2) + \text{cyc.}\right]. \end{aligned} \tag{3.27}$$

This becomes most obvious when we evaluate the reduced bispectrum, which at tree-level is given by

$$Q_{g,211} = \frac{1}{b_1}\,Q_{211} + \frac{b_2}{b_1^2} + 2\frac{\gamma_2}{b_1^2}\,\frac{K(\boldsymbol{k}_1,\boldsymbol{k}_2)\,P_1\,P_2 + \text{cyc.}}{P_1\,P_2 + \text{cyc.}}, \tag{3.28}$$

using the short-hand notation $P_i \equiv P_L(k_i)$. There are two crucial features to note about this expression: 1) each of the three terms has a different dependence on the wave vectors, allowing us to determine all three bias coefficients separately when fitting to data; and 2) Eq. (3.28) is independent of the amplitude of fluctuations, which means when combined with the power spectrum the bispectrum breaks the degeneracy between b_1 and σ_8 (Fry, 1994). This highlights one of the major strengths of the bispectrum, and we will encounter a similar example concerning the growth rate of structures in Sec. 3.3.3.

3.3 From theory to observations

A fundamental obstacle in measuring correlation functions of the density field is that we are only able to observe our own Universe, while the correlators are strictly defined as averages over various random realisations. Random fields for which it is possible to derive the same information from just a single realisation are called ergodic. Luckily, in our case this is a good assumption as the clustering within widely separated parts of the Universe should be causally unconnected. Following the cosmological principle, we can treat these patches effectively as independent outcomes of the same statistical process (Peebles, 1980), which allows us to replace the ensemble averages with averages

over volume. Yet, for this to hold the volume must be large compared to the correlation scale, so it can contain sufficiently many sub-volumes.

3.3.1 Simple power spectrum and bispectrum estimators

In this and the following section we consider a rather idealised survey geometry, namely a box of volume V with periodic boundaries, and filled with point-like objects of constant number density $\overline{n}$. Although too simplistic to match any realistic galaxy survey, this case serves to illustrate the main methodology, and in particular applies to the set-up of N-body simulations, which we make use of in later chapters. Due to the finite size of the survey, the Fourier transform yields a discrete set of modes, spaced out by the fundamental frequency $k_f = 2\pi/V^{1/3}$, which corresponds to the largest wavelength that fits into the box. A simple estimator (which we indicate by a hat) that follows from averaging over volume is[3] $\hat{P}(\boldsymbol{k}) = \langle|\delta n(\boldsymbol{k})|^2\rangle_\bullet/V - 1/\overline{n}$, where the components of $\boldsymbol{k}$ are multiples of k_f. In addition, we have subtracted the shot noise contribution (see Eq. 3.12), so that $\langle\hat{P}(\boldsymbol{k})\rangle = P(k)$. However, from statistical isotropy we expect that $\hat{P}$ should only depend on the magnitude of $\boldsymbol{k}$, so we can reduce noise by averaging over a spherical shell in Fourier space with thickness $\Delta k > k_f$. The estimator thus takes the form

$$\hat{P}(k) = \frac{1}{V_P(k)}\int_k \mathrm{d}^3q_1 \int_k \mathrm{d}^3q_2\, \delta_D(\boldsymbol{q}_{12})\, \frac{\langle\delta n(\boldsymbol{q}_1)\,\delta n(\boldsymbol{q}_2)\rangle_\bullet}{V} - \frac{1}{\overline{n}}\,, \tag{3.29}$$

where the integrals are evaluated on the interval $k-\Delta k/2 \le q_i \le k+\Delta k/2$ and $V_P(k)/k_f^3 \approx 4\pi\, k^2\,\Delta k/k_f^3$ denotes the number of modes per shell. For the bispectrum we can proceed in the same manner and the corresponding estimator reads

$$\begin{aligned}\hat{B}(k_1,k_2,k_3) = \frac{1}{V_B(k_1,k_2,k_3)}\int_{k_1} \mathrm{d}^3q_1 \int_{k_2} \mathrm{d}^3q_2 \int_{k_3} \mathrm{d}^3q_3\, \delta_D(\boldsymbol{q}_{123})\, \frac{\langle\delta n(\boldsymbol{q}_1)\,\delta n(\boldsymbol{q}_2)\,\delta n(\boldsymbol{q}_3)\rangle_\bullet}{V} \\ - \frac{1}{\overline{n}}\Big[\hat{P}(k_1)+\hat{P}(k_2)+\hat{P}(k_3)\Big] - \frac{1}{\overline{n}^2}\,,\end{aligned} \tag{3.30}$$

with the number of fundamental triangles in the three shells given by $V_B(k_1,k_2,k_3)/k_f^6 \approx 8\pi^2\, k_1 k_2 k_3\,(\Delta k)^3/k_f^6$. In Sec. 6.4.4 we will discuss an alternative to Eq. (3.30), which renders the three integrals separable and thus significantly speeds up the measurement process.

In a less idealised situation (Baumgart and Fry, 1991) showed that the above power spectrum and bispectrum estimators can still be constructed in a very similar way, provided that they account for the survey selection function. Furthermore, one can assign a weight to each galaxy, which can be chosen such that it maximises the signal-to-

[3] Once the point-like tracers of the continuous field have been identified, the average over the Poisson distribution has implicitly taken place. To reflect this, we include $\langle\ldots\rangle_\bullet$ in the estimators.

noise of the estimator (Feldman et al., 1994). These weights may reflect properties of the galaxies and their host dark matter halo and were recently shown to perform optimally when just depending on halo mass (Smith and Marian, 2015, 2016).

Finally, it is interesting to compare the total number of (mostly independent) configurations we can measure from the power spectrum and bispectrum, given a maximum wave number $k_{\rm max}$ and shell width Δk. For the former we simply add up all bins until $n_{\rm max} = k_{\rm max}/\Delta k$, whereas for the latter we show in Appendix 3.A that the number of triangles whose largest side is fixed to $k = i\,\Delta k$ scales as[4] $i(i+4)/4 - 1/8$, such that

$$N_{\rm conf.}\left(\hat{P}\right) = \sum_{i=1}^{n_{\rm max}} = n_{\rm max}\,, \tag{3.31}$$

$$N_{\rm conf.}\left(\hat{B}\right) = \sum_{i=1}^{n_{\rm max}} \left[\frac{i(i+4)}{4} - \frac{1}{8}\right] = \frac{1}{24}\left(2n_{\rm max}^3 + 15n_{\rm max}^2 + 10n_{\rm max}\right). \tag{3.32}$$

This demonstrates that the number of bispectrum configurations grows cubically with $k_{\rm max}$ and thus much faster than for the power spectrum, whose number of independent measurements only increases linearly.

3.3.2 Covariances and signal-to-noise

Since we aim to extract cosmological information from the correlation functions, it is vital that we are able to quantify the measurement error on each of the configurations, as well as the degree of their correlation. This is encoded in the covariance matrix C, e.g. for the power spectrum given by:

$$\mathsf{C}^P_{ij} \equiv \langle \hat{P}(k_i)\,\hat{P}(k_j)\rangle - \langle \hat{P}(k_i)\rangle\,\langle \hat{P}(k_j)\rangle\,. \tag{3.33}$$

By plugging in the estimator from Eq. (3.29) and expanding the ensemble averages using Wick's theorem, we can show that the covariance matrix splits into a Gaussian and non-Gaussian contribution (Meiksin and White, 1999; Scoccimarro et al., 1999b)

$$\mathsf{C}^P_{ij} = \frac{2(2\pi)^3}{V\,V_P(k_i)}\,\delta^K_{ij}\left(P(k_i) + \frac{1}{\overline{n}}\right)^2 + \overline{T}_{ij}\,, \tag{3.34}$$

where the latter is related to particular configurations of the trispectrum, i.e. the four-point correlation function of Fourier modes:

$$\overline{T}_{ij} = \frac{1}{V_P(k_i)\,V_P(k_j)}\int_{k_i} \mathrm{d}^3q_1 \int_{k_j} \mathrm{d}^3q_2\,\frac{\langle\langle \delta n(\boldsymbol{q}_1)\,\delta n(-\boldsymbol{q}_1)\,\delta n(\boldsymbol{q}_2)\,\delta n(-\boldsymbol{q}_2)\rangle_\bullet\rangle_c}{V^2}\,. \tag{3.35}$$

[4] Note that this number is about a factor of 2 smaller than what is reported e.g. in Sefusatti and Scoccimarro (2005).

We can identify a number of important features from Eq. (3.34). Firstly, even though we have subtracted shot noise from the estimator it contaminates the covariance matrix and thus increases the measurement errors (the same would be true if we were to consider the two-point correlation function ξ instead). Next, the covariance matrix scales inversely with survey volume, which is reasonable as a bigger volume effectively implies an average over a larger number of independent patches for a fixed scale k. And finally, as long as the Gaussian part of the covariance matrix dominates over $\overline{T}_{ij}$, power spectrum estimates at different scales are uncorrelated, which is not the case for ξ. The same features are shared by the bispectrum covariance matrix, which however picks up contributions from correlation functions up to sixth order, so in the following we only give the result in the Gaussian limit (see Sefusatti et al., 2006, for the full expression),

$$\begin{aligned} \mathsf{C}^B_{ij} &\equiv \langle \hat{B}(k_{i_1}, k_{i_2}, k_{i_3})\, \hat{B}(k_{j_1}, k_{j_2}, k_{j_3})\rangle - \langle \hat{B}(k_{i_1}, k_{i_2}, k_{i_3})\rangle \langle \hat{B}(k_{j_1}, k_{j_2}, k_{j_3})\rangle \\ &= \frac{s_B\,(2\pi)^6}{V\, V_B(k_{i_1}, k_{i_2}, k_{i_3})}\, \delta^K_{ij} \left(P(k_{i_1}) + \frac{1}{\overline{n}}\right)\left(P(k_{i_2}) + \frac{1}{\overline{n}}\right)\left(P(k_{i_3}) + \frac{1}{\overline{n}}\right) + \dots\,, \end{aligned} \tag{3.36}$$

where δ^K_{ij} is one if the two triangles defined by $\{k_{i_1}, k_{i_2}, k_{i_3}\}$ and $\{k_{j_1}, k_{j_2}, k_{j_3}\}$ are congruent, and zero otherwise. The symmetry factor $s_B = 6, 2, 1$ applies to equilateral, isosceles and scalene triangles, respectively.

Estimating the covariance matrices from the observed data itself is possible by splitting the survey volume into several sub-volumes and computing the intrinsic scatter with a jack-knife or bootstrap method. However, with increasing correlation scale the number of such sub-volumes drops rapidly, which limits the accuracy of the estimates and therefore makes them less and less reliable (Beutler et al., 2011; Norberg et al., 2009). For that reason the covariance matrices are usually predicted, either theoretically from the expressions in Eqs. (3.34) and (3.36), or based on numerical simulations. As we have seen previously, density perturbations develop non-Gaussianities, which implies that the higher-order contributions to the covariance matrices will become important. In that case the theoretical computation is very involved, which is why the current default is to estimate covariance matrices from an independent sample of N-body simulations with varying initial conditions (see Monaco, 2016, for a recent review). To get a reliable estimate the sample size has to be larger than the dimension of C by at least a factor of a few (Hartlap et al., 2007). Assuming a survey volume of $8\,(\mathrm{Gpc}/h)^3$ and $k_{\mathrm{max}} = 0.2\,\mathrm{Mpc}/h$ with $\Delta k = 2k_f$, the number of measurable power spectrum configurations is about $N_{\mathrm{conf.}}\left(\hat{P}\right) \approx 30$, whereas $N_{\mathrm{conf.}}\left(\hat{B}\right) \approx 3000$ for the bispectrum, according to Eqs. (3.31) and (3.32). Generating a large enough sample required to obtain a reliable bispectrum covariance matrix is consequently very costly and therefore presents one of the major challenges for a bispectrum measurement. This is the main motivation for the work

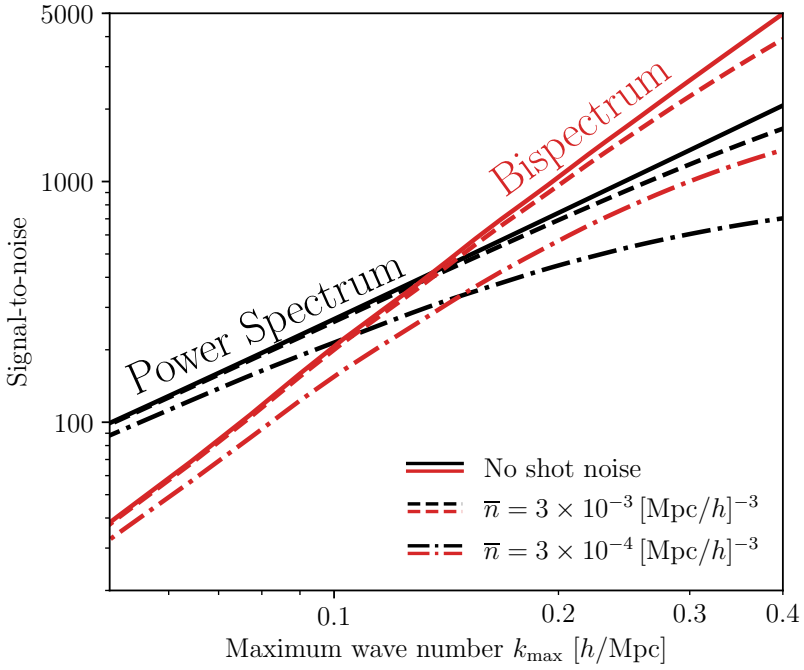


Fig. 3.5 Comparison of the cumulative signal-to-noise for the galaxy power spectrum (blue) and bispectrum (red) as a function of the maximum wave number k_{max}. Different line styles correspond to different number densities of tracers. For this plot we have assumed a survey volume of 8 $(\mathrm{Gpc}/h)^3$, and lowest order perturbation theory for P_g and B_g, as well as their covariance matrices.

presented in Chap. 6, where we are looking for ways to shrink the size of the covariance matrix.

How much can we gain from measuring higher-order statistics? Based on their covariances matrices, we can evaluate the signal-to-noise of the power spectrum and bispectrum estimators, which is a good proxy for their constraining power:

$$\left(\frac{S}{N}\right)^2_P = \sum_{i,j=1} \langle \hat{P}_i \rangle \left(\mathsf{C}^P_{ij}\right)^{-1} \langle \hat{P}_j \rangle\,, \quad \left(\frac{S}{N}\right)^2_B = \sum_{i,j=1} \langle \hat{B}_i \rangle \left(\mathsf{C}^B_{ij}\right)^{-1} \langle \hat{B}_j \rangle\,, \tag{3.37}$$

where $\hat{P}_i \equiv \hat{P}(k_i)$ and $\hat{B}_i \equiv \hat{B}(k_{i_1}, k_{i_2}, k_{i_3})$. Both sums run over all configurations and thus make the signal-to-noise a function of the maximum wave number k_{max}. These expressions are straightforward to evaluate if we assume Gaussian covariance matrices and apply leading order perturbation theory to predict the clustering signals for a given distribution of galaxies, i.e. Eqs. (3.24) and (3.27). Adopting again a survey volume of 8 $(\mathrm{Gpc}/h)^3$, and using the linear power spectrum corresponding to a ΛCDM model with values given in Table 1.1 and the following bias parameters[5],

$$b_1 = 1.5\,, \qquad b_2 = -0.69\,, \qquad \gamma_2 = -0.14\,, \tag{3.38}$$

[5]The choice of b_2 is motivated by the peak-background split prediction from Lazeyras et al. (2016), while for γ_2 we use the local Lagrangian value, i.e. $\gamma_2 = -2/7\,(b_1 - 1)$ (Chan et al., 2012).

we obtain the plot shown in Fig. 3.5. The solid lines correspond to a case with negligible shot noise, whereas the other two distinguish different tracer number densities. The figure illustrates that the bispectrum signal-to-noise grows more quickly with maximum wave number than the power spectrum and comes to dominate at $k_{\mathrm{max}} \sim 0.15\,\mathrm{Mpc}/h$. Overall we can therefore expect the bispectrum to yield valuable additional information. On the other hand, the plot also shows that the bispectrum is slightly more impacted by shot noise, which tends to suppress the signal-to-noise with decreasing $\overline{n}$ — a trend that continues with statistics of even higher orders (Sefusatti and Scoccimarro, 2005).

3.3.3 Observational effects

A galaxy survey measures the redshifts and angular locations of galaxies, and as such cannot directly determine their three-dimensional positions. Applying the Hubble expansion law allows us to convert redshifts into distances, but requires assumptions about the underlying cosmological model and does not account for contributions from the peculiar velocities of the galaxies. Consequently, this conversion step introduces systematic errors in the distance estimates, which are — remarkably — not just a nuisance, but can also be an important source of information.

Redshift space distortions

Because peculiar velocities make up part of the observed redshift, we infer comoving galaxy positions $\boldsymbol{s}$ that are shifted away from the true locations $\boldsymbol{x}$ in radial, or line-of-sight, direction:

$$\boldsymbol{s} = \boldsymbol{x} - f\, u_r(\boldsymbol{x})\, \hat{\boldsymbol{x}}\,, \tag{3.39}$$

where $\hat{\boldsymbol{x}} = \boldsymbol{x}/x$ and u_r is the radial component of the scaled peculiar velocity field $\boldsymbol{u}(\boldsymbol{x}) \equiv -v(\boldsymbol{x})/(\mathscr{H} f)$. This leads to observable effects on the clustering patterns of galaxies, known as *redshift space distortions*, which display a qualitatively different behaviour on large and small scales. On large scales gravitational instability generates bulk flows in the direction of density gradients (see Sec. 3.2.1), so within large spherical over-densities, as depicted in the left-hand panel of Fig. 3.6, matter tends to move towards the centres. That makes galaxies positioned at the front of the over-density (with respect to the observer) recede faster than the Hubble flow, increasing their redshifts, whereas galaxies on the far-edge move slower, which results in them having smaller redshifts. Therefore, we over-, or underestimate their true distances, respectively, which lets the large-scale over-densities appear squashed in the line-of-sight direction — the *Kaiser effect* (Kaiser, 1987). On the other hand, shell crossing causes a virialisation of the small-scale velocity field, such that structures appear elongated, as shown based on

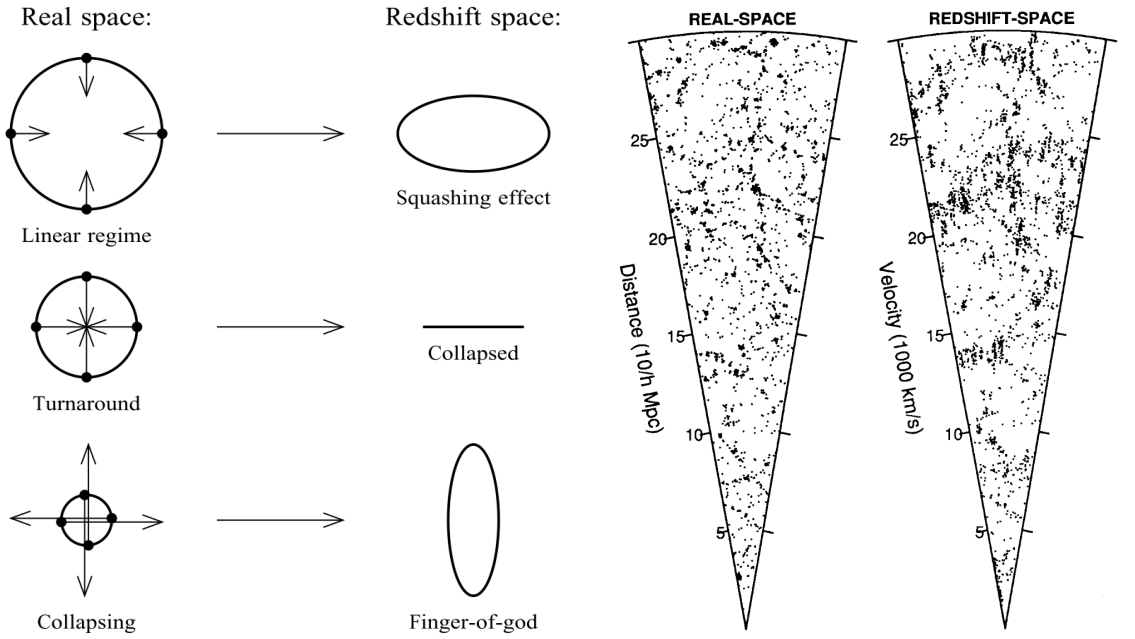


Fig. 3.6 *Left:* Redshift space distortions, demonstrated based on the collapse of a spherically symmetric over-density for various scales (Hamilton, 1998, from). *Right:* Wedge from an N-body simulation, displaying the positions of particles in real and redshift space (from Praton et al., 1997). In both cases the observer is imagined to be at the bottom of the illustration.

the example of a collapsing spherical over-density in Fig. 3.6. Because these structures seem to be pointing at the observer, this is commonly referred to as *fingers-of-god* effect (Jackson, 1972). This can be clearly seen in the right-hand panel of Fig. 3.6, which compares a real and redshift space wedge extracted from an N-body simulation.

The redshift space squashing leads to an increase of the clustering amplitude along the line-of-sight, while the small-scale stretching has the opposite effect. The former can be computed perturbatively in the linear regime, and in the distant-observer approximation where the line-of-sight is considered to be parallel for all galaxies, it gives rise to the factor $1+\beta\mu^2$ with $\beta \equiv f/b_1$ and μ being the cosine of the angle between wave vector $\boldsymbol{k}$ and the line-of-sight (Kaiser, 1987; Scoccimarro et al., 1999a). The fingers-of-god effect is often accounted for by a (phenomenological) Lorentzian damping term (Davis and Peebles, 1983; Park et al., 1994; Peacock and Dodds, 1994) depending on the pairwise velocity dispersion σ_p, so that the redshift space power spectrum can be written as

$$P_{g,s}(k,\mu) = \left(1+\beta\mu^2\right)^2 \frac{P_g(k)}{1+k^2\mu^2\sigma_p^2/2}\,. \tag{3.40}$$

This implies that a measurement of $P_{g,s}$ as a function of μ has the potential to constrain the growth rate f, giving us an opportunity to test general relativity which predicts that $f \approx \Omega_m^{5/9}$ (Bouchet et al., 1995). However, due to the degeneracy between b_1 and σ_8, it is

only possible to extract the parameter β from the multipoles of $P_{g,s}$ (Hamilton, 1992), but not the growth rate on its own. This situation changes when including the redshift space bispectrum, since redshift space distortions induce a non-trivial dependence on wave vectors (Hivon et al., 1995; Scoccimarro et al., 1999a), similar to the effect of higher-order bias terms (Sec. 3.2.2). By measuring bispectrum multipoles for various triangle configurations, it is therefore possible to break the degeneracy between f and b_1, which can improve constraints on the growth rate by a factor of 2 to 3 (Gagrani and Samushia, 2017; Song et al., 2015).

Alcock-Paczyński effect

A different kind of distortion effect is introduced by our assumption of a cosmological model in order to convert redshifts and angular separations into proper distances. If we consider a pair of galaxies with redshifts $z \pm \Delta z$, separated by the angle $\Delta\theta$ on the sky, their radial and transverse separations are given by

$$\Delta r_\parallel = \frac{\Delta z}{H(z)}\,, \qquad \Delta r_\perp = (1+z)\, D_A(z)\, \Delta\theta\,, \tag{3.41}$$

and thus sensitive to the Hubble rate and angular diameter distance $D_A(z)$ at the mean redshift z. Assuming an incorrect set of cosmological parameters leads to a squashing or stretching in line-of-sight and transverse direction, encoded by the factors

$$f_\parallel = \frac{\Delta r_\parallel}{\Delta r_{\parallel,a}} = \frac{H_a(z)}{H(z)}\,, \qquad f_\perp = \frac{\Delta r_\perp}{\Delta r_{\perp,a}} = \frac{D_A(z)}{D_{A,a}(z)}\,, \tag{3.42}$$

where the subscript "a" denotes quantities derived from the incorrect cosmology. Having a distribution of objects that is known to be intrinsically spherical, Charles Alcock and Bohdan Paczyński suggested to use these distortions as a cosmological test — henceforth, the Alcock-Paczyński (AP) test — as in this case we can require that $f_\parallel / f_\perp = 1$, allowing us to constrain the combination $H(z)\, D_A(z)$ (Alcock and Paczynski, 1979). We can conduct this AP test based on the clustering of galaxies, since statistical isotropy dictates that the clustering properties should be identical in all directions. However, as we have seen above, redshift space distortions also produces a squashing as well as stretching, which must be carefully modelled in order to be distinguishable from the AP signal (Ballinger et al., 1996; Matsubara and Suto, 1996).

If we additionally know the true size of some physical scale, it is possible to constrain $f_\parallel$ and $f_\perp$ separately, which breaks the degeneracy between $H(z)$ and $D_A(z)$. This is the case for the baryon acoustic oscillations (BAO), which imprint a characteristic feature in the galaxy correlation functions, whose scale can be accurately determined from CMB experiments. By measuring its scale at various redshifts we can reconstruct the expansion history of our Universe, which is a powerful probe of dark energy (Eisenstein

et al., 1998; Glazebrook and Blake, 2005; Weinberg et al., 2013). Moreover, the BAO feature resides on relatively large scales ($\sim 100\,\mathrm{Mpc}/h$), making it robust under the influence of non-linear gravitational evolution (though see Crocce and Scoccimarro, 2008; Smith et al., 2008), which is why it has become one of the main pillars in the analysis of galaxy surveys.

3.3.4 Galaxy redshift surveys: from past till future

While measurements of the two-point correlation function and power spectrum are now routinely carried out from large galaxy catalogues, they rely on a vast amount of previous efforts and studies. The earliest of these were limited to angular galaxy correlations, which were extracted by Jim Peebles and collaborators (e.g. Hauser and Peebles, 1973; Peebles and Hauser, 1974; Yu and Peebles, 1969) who considered two-dimensional maps compiled by Donald Shane and Carl Wirtanen from thousands of individual photographic plates (the *Lick catalogue*) (Shane and Wirtanen, 1967). Three-dimensional analyses, as discussed in this chapter, only became possible with the advent of redshift surveys, and the first of these — the *CfA Redshift Survey* (Huchra et al., 1983) — recorded redshifts of about 2400 galaxies between 1977 and 1982, giving rise to the correlation function results of Davis and Peebles (1983). From a present-day perspective this number of galaxies appears rather modest because the extent of redshift surveys increased dramatically with the development of fibre-optic and multi-slit spectrographs in the 1990s, which allowed to take spectra of several hundred galaxies at the same time. Equipped with this technology the *Two-degree-Field (2dF) Galaxy Redshift Survey* (Colless et al., 2001) already contained positions of about 220,000 galaxies, which enabled the first precise measurements of the power spectrum (Percival et al., 2001). Another major leap forward was undertaken by the *Sloan Digital Sky Survey (SDSS)* (York et al., 2000) with over a million galaxies that provided enough statistical power to detect the BAO feature in the two-point correlation function (Eisenstein et al., 2005) and in the same year in the power spectrum from *2dF* (Cole et al., 2005). State-of-the-art measurements derive from the *Baryon Oscillation Spectroscopic Survey (BOSS)* (Dawson et al., 2013), which completed mapping 1.5 million galaxies to redshift $z = 0.7$ in 2014. The resulting two-point clustering analyses, along with the BAO technique described in Sec. 3.3.3, have produced stringent constraints on the parameters of the ΛCDM model and its extensions (Alam et al., 2017) (see also Secs. 1.1 and 1.2).

The three-point correlation function and bispectrum have been studied far less in comparison, mainly due to the obstacles related to their modelling and the estimation of their covariance matrices as outlined in Sec. 3.2 and 3.3.2. In addition, early studies (Fry and Seldner, 1982; Groth and Peebles, 1977; Peebles and Groth, 1975) from the *Lick catalogue* were limited to small scales, making their interpretation with perturbation theory

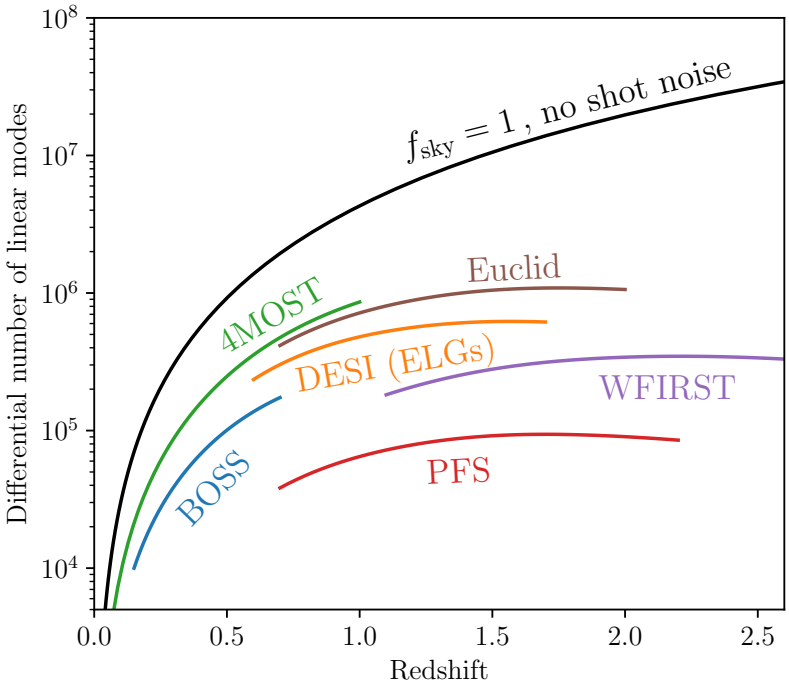


Fig. 3.7 Differential number of modes, $\mathrm{d}N/\mathrm{d}z$, which can be used in an analysis based on perturbation theory. The maximally achievable number (full sky, no shot noise) is shown in black, whereas the coloured lines represent various redshift surveys, whose characteristics are summarised in Table 3.1. Results are derived from a ΛCDM model with parameters close to those in Table 1.1.

rather difficult. For that reason the configuration dependence predicted by Eq. (3.22) was only confirmed by Frieman and Gaztañaga (1999) using large-scale data from the *APM survey* (Maddox et al., 1990). More extensive measurements were finally reported for the *IRAS Point Source Catalogue* (Scoccimarro et al., 2001) and the *2dF Survey* (Verde et al., 2002), which explicitly demonstrated the potential of the bispectrum for breaking the degeneracies between the linear bias parameter, amplitude of density fluctuations and growth rate. The most recent studies were carried out for *BOSS* and one of which detected the BAO in the three-point function (Slepian et al., 2017b), whereas another combined the power spectrum and bispectrum for the first time in a joint analysis (Gil-Marín et al., 2015, 2017), finding slight tension with the growth rate prediction from general relativity. However, we note that the latter analysis was based on a bispectrum model that ignored contributions from higher-order perturbation theory, which can lead to inconsistencies with the power spectrum and motivates our work in Chap. 7.

The next decade will see the implementation of a multitude of new redshift surveys, which will bring about significant improvements over all previous measurements, because their capabilities exceed those of *BOSS* in one or several ways, including redshift range, survey depth and sky coverage. Future surveys are often compared in terms of their total effective volume, meaning the survey volume that is well sampled with galaxies and not dominated by shot noise. This is a good proxy for how accurate the

Table 3.1 Redshift range, observed number of objects, comoving number density and sky coverage for a range of forthcoming galaxy redshift surveys — *DESI (ELGs)*, *PFS*, *4MOST*, *WFIRST*, and *Euclid* (Amendola et al., 2018; de Jong et al., 2016; DESI Collaboration et al., 2016; Spergel et al., 2015; Takada et al., 2014, respectively) — compared with *BOSS* (Reid et al., 2016).

Survey	Δz	Number of galaxies	$\overline{n}$ $[(h/\mathrm{Mpc})^3]$	$\Delta\Omega$ [deg^2]
BOSS	$0.15-0.7$	1.4×10^6	2.5×10^{-4}	9,376
DESI (ELGs)	$0.6-1.7$	17×10^6	3.2×10^{-4}	14,000
PFS	$0.7-2.2$	4×10^6	4.6×10^{-4}	1,464
4MOST	$0.05-1$	15×10^6	7.3×10^{-4}	15,000
WFIRST	$1.1-2.6$	18×10^6	12×10^{-4}	2,200
Euclid	$0.7-2$	40×10^6	5.2×10^{-4}	15,000

BAO scale can be determined, but when aiming to extract cosmological information by fitting a perturbation theory motivated power spectrum or bispectrum model to the data, it is more useful to compare the total number of modes that can be utilised for this purpose. That allows us to take into account that the range of scales where perturbation theory is applicable increases with redshift, as we have seen in Sec. 3.2.1. Using that there are $4\pi k^2\Delta k/k_f^3$ modes within a spherical shell of width Δk, the total number of modes covered by a survey with redshift range $z\pm\Delta z$ ($\Delta z/z\ll1$) and subtending the solid angle $\Delta\Omega=4\pi f_{\mathrm{sky}}$ is given by

$$\Delta N=\frac{\Delta\Omega\,\Delta z}{2\pi^2}\frac{\mathrm{d}^2V_c(z)}{\mathrm{d}\Omega\,\mathrm{d}z}\int_0^{k_\star(z)}\mathrm{d}k\,k^2\left[\frac{\overline{n}P_g(k,z)}{1+\overline{n}P_g(k,z)}\right]^2, \tag{3.43}$$

where the term in the square brackets accounts for a shot noise weighting of each mode (Tegmark, 1997), and $\mathrm{d}^2V_C(z)/\mathrm{d}\Omega\mathrm{d}z$ stands for the differential comoving volume. Assuming that the regime of validity for perturbation theory scales as $k_\star(z)=0.1\left[D(z=0)/D(z)\right]h/\mathrm{Mpc}$ (Crocce and Scoccimarro, 2006a), Fig. 3.7 shows the differential number of modes, $\mathrm{d}N/\mathrm{d}z$, for *BOSS* and a selection of upcoming surveys (details are given in Table 3.1) as a function of redshift. We also compare to the full-sky case with negligible shot noise (black line), and we see that the combination of future surveys comprise a large fraction of the available modes. As the signal-to-noise roughly scales as the square root of the total number of modes, i.e. the area underneath each of the curves, the plot clearly demonstrates the potential of large-scale structure studies for the next ten years. For instance, *DESI* and *WFIRST* obtain about 10 times as many modes as *BOSS*, whereas *Euclid* extends this by another factor of two.

Appendices

Appendix 3.A Number of triangle configurations

In this appendix we will derive an exact polynomial expression for the number of distinct triangle configurations of the bispectrum, given a maximum wave number k_{max} and shell width Δk. In general, the bispectrum is a function of the three wave vectors $\boldsymbol{k}_1$, $\boldsymbol{k}_2$ and $\boldsymbol{k}_3$, which under the assumption of statistical homogeneity must form a closed triangle, such that $\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3 = 0$. Let us further assume statistical isotropy, in which case the bispectrum is completely described in terms of the magnitudes of the three wave vectors. Because of its symmetry it is sufficient to consider $k_1 \geq k_2 \geq k_3$, and due to the closure condition we further have the inequality $k_1 \leq k_2 + k_3$. These two relations are visualised as the shaded triangles (light green and blue) in Fig. 3.A.1, which shows k_2 on the x- and k_3 on the y-axis, both up to a maximum value of k_1. The allowed triangle configurations are thus given by the overlap, i.e. the smaller triangle that appears in dark green.

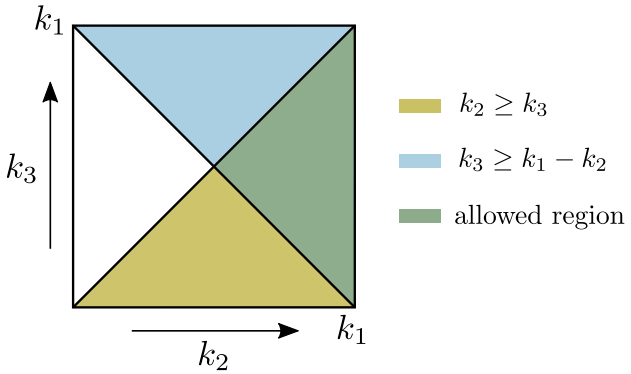


Fig. 3.A.1 Pairs of (k_2, k_3) for a fixed value of k_1. Distinct configurations are constrained to the light green triangular region, whereas the closure condition only allows (k_2, k_3) to take values in the blue triangular region. Their overlap is shown in dark green.

We now take the three wave numbers to be integer multiples of the shell width: $k_1 = n\,\Delta k$, $k_2 = i\,\Delta k$, and $k_3 = j\,\Delta k$. For a fixed value of n, we then have to add up all

bins contained in the dark green triangle. If n is even, we get:

$$N_{\mathrm{even}}(n) = \sum_{i=n/2}^{n} \sum_{j=\max\{1,n-i\}}^{i} = -1 + \sum_{i=n/2}^{n} \sum_{j=n-i}^{i} = -1 + \sum_{i=n/2}^{n} \Big[1+2i-n\Big] = \frac{1}{4}(2+n)^2 - 1\,, \tag{3.44}$$

where in the second step we have used that $j = 0$ can only occur for the single bin where $i = n$, and the final sum was evaluated using Faulhaber's formula. Similarly, for n odd we find:

$$N_{\mathrm{odd}}(n) = \sum_{i=(n+1)/2}^{n} \sum_{j=\max\{1,n-i\}}^{i} = \frac{1}{4}(3+4n+n^2) - 1\,. \tag{3.45}$$

The total number of triangle configurations up to the scale $k_{\max} = n_{\max}\Delta k$ is finally given by the summation over Eqs. (3.44) and (3.45). Distinguishing between an even and odd value of $n_{\max}$ leads to

$$\begin{aligned} N_{\mathrm{conf.}}(n_{\max}) &= \begin{cases} \sum_{n=1}^{n_{\max}/2} N_{\mathrm{even}}(2n) + \sum_{n=1}^{n_{\max}/2} N_{\mathrm{odd}}(2n-1)\,, & n_{\max} \quad \text{even} \\ \sum_{n=1}^{(n_{\max}-1)/2} N_{\mathrm{even}}(2n) + \sum_{n=1}^{(n_{\max}+1)/2} N_{\mathrm{odd}}(2n-1)\,, & n_{\max} \quad \text{odd} \end{cases} \\ &= \begin{cases} \frac{n_{\max}}{24}\left(10+15n_{\max}+2n_{\max}^2\right)\,, & n_{\max} \quad \text{even} \\ \frac{n_{\max}}{24}\left(10+15n_{\max}+2n_{\max}^2\right) - \frac{1}{8}\,, & n_{\max} \quad \text{odd} \end{cases} \end{aligned} \tag{3.46}$$

where we have again made use of Faulhaber's formula. This demonstrates that the number of bispectrum configurations grows cubically with the maximum wave number. Ignoring the little difference between even and odd $n_{\max}$, it is also easy to show that

$$N_{\mathrm{conf.}}(n_{\max}) = \frac{1}{4}\sum_{n=1}^{n_{\max}} \left[i(i+4) - \frac{1}{2} \right] = \frac{n_{\max}}{24}\left(10+15n_{\max}+2n_{\max}^2\right)\,, \tag{3.47}$$

which is correct up to negligible rounding errors. Compared to Sefusatti and Scoccimarro (2005), who use the approximate formula $N_{\mathrm{conf.}} = \sum_{n=1}^{n_{\max}} n(n+1)/2$, Eq. (3.47) is smaller by about a factor of two.

Chapter 4

The anisotropic line correlation function as a probe of anisotropies in galaxy surveys

At the end of Sec. 3.1.3 we have seen how gravitational evolution induces correlations in the phases of Fourier modes. In this chapter we introduce the line correlation function — a statistical quantity designed to measure these phase correlations and therefore highly complementary to the two-point function and power spectrum. We propose an anisotropic generalisation of the line correlation function, which allows us to detect both, the growth rate of structures from redshift space distortions, as well as the AP effect through a squashing or stretching in line-of-sight and transverse directions (see Sec. 3.3.3). We demonstrate our formalism and compare to the two-point function based on two-dimensional mock density fields and the Zel'dovich approximation (Sec. 2.2.4).

4.1 Introduction

Galaxy surveys map a growing fraction of our Universe at ever increasing accuracy and are routinely used to test cosmological models. So far, agreement with the Λ cold dark matter (ΛCDM) model, a spatially flat universe primarily filled with dark matter and dark energy in the form of a cosmological constant, is excellent. Among their key science questions are the nature of dark matter, dark energy and the concrete model of the early Universe during which primordial density fluctuations were generated. A range of next-generation experiments, such as *DESI* (DESI Collaboration et al., 2016), *Euclid* (Amendola et al., 2018) and *SKA* (Dewdney et al., 2009), etc. are designed to shed light on to these questions.

To make progress, it is crucial to optimise the amount of information extracted from the raw data that these experiments produce. Up until recently, the majority of the

literature focused on two-point statistics, i.e. the power spectrum P or the two-point function ξ, to analyse the observed distribution of galaxies. If the density field were Gaussian, two-point statistics would be sufficient to specify the statistical properties of matter fluctuations. However, during the non-linear stage of gravitational interaction, where structures like filaments, clusters and eventually galaxies form, any initial density field becomes highly non-Gaussian. Hence, by focusing on two-point statistics, a considerable part of the available information is ignored.

Consequently, new statistical measures have received significant attention. Besides computing higher-order correlations (e.g. Frieman and Gaztanaga, 1994; Peebles, 1980; Peebles and Groth, 1975; Scoccimarro et al., 1998b), it was suggested to employ genus statistics (Gott et al., 1986; Hikage et al., 2002; Hoyle et al., 2002), Minkowski functionals (Hikage et al., 2003; Mecke et al., 1994) and other measures. However, all of these measures suffer from the conceptual limitation that they are correlated with the two-point function. Thus, an intriguing idea is to exclusively analyse information not already contained in two-point statistics, i.e. phase information.

Correlations in the phases of Fourier coefficients emerge as a consequence of non-Gaussianity; since two-point statistics depend on the amplitudes of Fourier coefficients only, they are blind to phase information. The phase factors of the matter density field have been studied many years ago, but much of the original work in Jain and Bertschinger (1998); Ryden and Gramann (1991); Soda and Suto (1992) focused on the evolution of single phases away from their initial values. A first statistic that described the phase difference between neighbouring Fourier modes was presented in Chiang (2001); Coles and Chiang (2000); Watts et al. (2003). Shortly thereafter, their considerations were generalised by Matsubara (2003), who quantified phase information by computing the joint probability density of phase factors, which illuminated its relation to higher-order spectra (see also Hikage et al., 2004; Matsubara, 2007)). In Hikage et al. (2005), the probability densities were applied to a galaxy catalogue from the Sloan Digital Sky Survey, Data Release Two, giving constraints on bias models. Recently, a new measure for phase information, the line correlation function, was introduced by Obreschkow et al. (2013) as a spherically-averaged three-point correlation of phase factors. In the follow-up paper by Wolstenhulme et al. (2015), it was computed perturbatively and related to the bispectrum and power spectrum at leading order.

The aim of this chapter is to extend the line correlation function to account for line-of-sight and transverse distance scales separately. The introduction of such an anisotropic line correlation function (ALCF) opens up the possibility to detect anisotropies and distortions along the line-of-sight, which are inherently present in any galaxy survey due to redshift space distortions.

This chapter is organised as follows: after a brief review of the definition and properties of the line correlation function in Sec. 4.2; we present and test a modification in

Sec. 4.3, which is capable of quantifying anisotropies. In Sec. 4.4 this new formalism is applied to simple mock fields and we investigate the sensitivity to the Alcock-Paczyński (AP) effect and kinematical redshift-space distortions in comparison to results based on the two-point function. We conclude in Sec. 4.5.

4.2 The isotropic line correlation function

In this section, we review the isotropic line correlation function (ILCF), which serves as the starting point for our generalisation to the ALCF in Sec. 4.3.

4.2.1 Definition

Given a density field $\delta(\boldsymbol{x})$, the ILCF measures correlations in its Fourier phase factors which are obtained by dividing amplitudes in Fourier space[1],

$$\delta(\boldsymbol{x}) \xrightarrow{\text{FT}} \delta_{\boldsymbol{k}} \xrightarrow{\text{whitening}} \epsilon_{\boldsymbol{k}} \equiv \frac{\delta_{\boldsymbol{k}}}{|\delta_{\boldsymbol{k}}|} \xrightarrow{\text{IFT}} \epsilon(\boldsymbol{x})\,. \tag{4.1}$$

The *whitened* density field $\epsilon(\boldsymbol{x})$ is thus devoid of any information probed by two-point statistics. Consequently, the simplest measure of phase information must be based on the three-point correlator and since Obreschkow et al. (2013) observed (phenomenologically) that the whitening transformation tends to collapse elongated objects to thin line segments, they chose these three points to be distributed equidistantly on a straight line. More precisely, and expressed in Fourier space, the ILCF is defined by

$$\ell(r) \equiv \frac{V^3}{(2\pi)^{3D}} \left(\frac{r^D}{V}\right)^{3/2} \iiint\limits_{\substack{|\boldsymbol{k}_1|,\,|\boldsymbol{k}_2|,\,|\boldsymbol{k}_3| \\ \leq 2\pi/r}} \mathrm{d}^D k_1\, \mathrm{d}^D k_2\, \mathrm{d}^D k_3\, \mathrm{e}^{i[\boldsymbol{k}_1\cdot\boldsymbol{x}+\boldsymbol{k}_2\cdot(\boldsymbol{x}+\boldsymbol{r})+\boldsymbol{k}_3\cdot(\boldsymbol{x}-\boldsymbol{r})]}\, \langle \epsilon_{\boldsymbol{k}_1}\, \epsilon_{\boldsymbol{k}_2}\, \epsilon_{\boldsymbol{k}_3} \rangle\,, \tag{4.2}$$

where D denotes the dimension of space, V the volume of the survey and $\langle\rangle$ an ensemble average over all realisations of the density field. In writing Eq. (4.2) we adopted the same conventions as employed in Wolstenhulme et al. (2015).

The ILCF differs from the conventional three-point correlator of whitened density fields in two respects. Firstly, it is multiplied by a prefactor $(r^D/V)^{3/2}$ and secondly, each field is convolved with a top-hat window function that cuts off high-frequency modes $|\boldsymbol{k}| > 2\pi/r$. These corrections are introduced to regularise the integral in the regimes of large and small scales, respectively. As discussed in Obreschkow et al. (2013), Fourier modes which are associated with scales larger than the longest physical correlation length or scales smaller than the lowest characteristic length obtain random phase factors that would cause the ILCF to be dominated by random noise.

[1] The zero mode is conventionally set to zero, i.e. $\epsilon_{\boldsymbol{k}=0} = 0$.

It is important to reiterate that the ILCF is independent of amplitude information and for that reason not just a regularised integral of the bispectrum. It depends on the quantity $\langle \epsilon_{\boldsymbol{k}_1} \epsilon_{\boldsymbol{k}_2} \epsilon_{\boldsymbol{k}_3} \rangle = \langle \exp\left[i \left(\theta_{\boldsymbol{k}_1} + \theta_{\boldsymbol{k}_2} + \theta_{\boldsymbol{k}_3} \right) \right] \rangle$, which can be determined by using the probability density function (PDF) of Fourier phases, $\mathscr{P}\{\theta\}$. This PDF was computed perturbatively for mildly non-Gaussian fields by Matsubara (2003), showing that it is not solely related to the bispectrum but a progression of all higher order spectra such as the trispectrum etc. Taking this result, Wolstenhulme et al. (2015) showed that the ILCF can be expressed to lowest order as follows:

$$\ell(r) \simeq \left(\frac{\sqrt{\pi}}{2}\right)^3 \left(\frac{r}{2\pi}\right)^{3D/2} \iint\limits_{\substack{|\boldsymbol{k}_1|,\,|\boldsymbol{k}_2|,\\ |\boldsymbol{k}_1+\boldsymbol{k}_2| \leq 2\pi/r}} \mathrm{d}^D k_1\, \mathrm{d}^D k_2\, \frac{B(\boldsymbol{k}_1, \boldsymbol{k}_2, -\boldsymbol{k}_1 - \boldsymbol{k}_2)}{\sqrt{P(k_1)\, P(k_2)\, P(|\boldsymbol{k}_1 + \boldsymbol{k}_2|)}} \times \cos\left[(\boldsymbol{k}_1 - \boldsymbol{k}_2) \cdot \boldsymbol{r} \right], \tag{4.3}$$

where $P(k)$ and $B(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$ are the ordinary power spectrum and bispectrum. Accordingly, we can think of the ILCF as compressing information from all higher order spectra into a single function. Only in the case of weakly non-Gaussian fields does the dominant contribution stem from the bispectrum.

For evaluating the ILCF on a given density field, Eq. (4.3) is not particularly useful. However, we arrive at a simple prescription to extract the line correlation by assuming ergodicity and replacing the ensemble average in Eq. (4.2) by an average over all translations and rotations. Integrating out the translation vector $\boldsymbol{x}$ as well as the orientation of $\boldsymbol{r}$ thus gives

$$\ell(r) = \frac{V^2}{(2\pi)^{2D}} \left(\frac{r^D}{V}\right)^{3/2} \iint\limits_{\substack{|\boldsymbol{k}_1|,\,|\boldsymbol{k}_2|,\\ |\boldsymbol{k}_1+\boldsymbol{k}_2| \leq 2\pi/r}} \mathrm{d}^D k_1\, \mathrm{d}^D k_2\, w_D\left(|\boldsymbol{k}_1 - \boldsymbol{k}_2|\, r\right) \frac{\delta_{\boldsymbol{k}_1}\, \delta_{\boldsymbol{k}_2}\, \delta_{-\boldsymbol{k}_1-\boldsymbol{k}_2}}{|\delta_{\boldsymbol{k}_1}\, \delta_{\boldsymbol{k}_2}\, \delta_{-\boldsymbol{k}_1-\boldsymbol{k}_2}|}, \tag{4.4}$$

where we have defined the rotational average of the exponential Fourier factor as

$$w_D(k\,r) \equiv \langle \mathrm{e}^{i \boldsymbol{k}\cdot\boldsymbol{r}} \rangle_{\mathscr{R}} = \begin{cases} \cos(k\,r)\,, & \text{if } D = 1\,, \\ J_0(k\,r)\,, & \text{if } D = 2\,, \\ \frac{\sin(k\,r)}{k\,r}\,, & \text{if } D = 3\,. \end{cases} \tag{4.5}$$

A discretised version of Eq. (4.4) will be used for all numerical computations throughout this chapter. Details on the implementation are given in App. 4.A.

4.2.2 Properties and applications

The morphology of the cosmic web, composed of structures such as clusters, filaments and voids, is largely encoded in the phases of Fourier coefficients (for a demonstration see Chiang and Coles, 2000; Coles and Chiang, 2000). Against this background, it is

natural to assume that the ILCF is sensitive to these kinds of objects and, owing to its linear configuration, especially to filaments. In fact, after a more systematic analysis, Obreschkow et al. (2013) conclude that the ILCF is generally a measure for aspherical structures on scales $\sim 2r$. This holds true for oblate and prolate objects alike, so that $\ell(r)$ is able to probe both cosmic sheets and filaments.

Furthermore, they find two scaling relations depending on the size and density of the substructure. Letting r_0 denote its characteristic scale and N the number of objects[2], their filling factor is given by $f = N V_0 / V$ where V_0 is the volume of a D-dimensional sphere of radius r_0. A change of the characteristic scale or the number of objects has the following effect on the ILCF:

$$a) \quad \left.\begin{array}{l} r_0 \to \alpha\, r_0 \\ f \ = \text{const.} \end{array}\right\} \quad \Rightarrow \quad \ell(r) \to \ell(r/\alpha)\,, \tag{4.6}$$

$$b) \quad N \to \alpha\, N \quad \Rightarrow \quad \ell(r) \to \frac{\ell(r)}{\sqrt{\alpha}}\,. \tag{4.7}$$

The first relation specifically targets the behaviour of $\ell(r)$ under a re-scaling of the substructure and therefore should not be confused with a scaling law resulting from power law clustering. Such a scaling law can be derived from Eq. 4.3 and assuming the hierarchical model (e.g. Bernardeau et al., 2002), so that we have $B \sim P^2$. For a power law spectrum $P(k) \sim k^n$ we then get

$$\ell(\alpha\, r) = \alpha^{-\frac{D+n}{2}}\, \ell(r)\,, \tag{4.8}$$

implying that the slope of $\ell(r)$ is less steep than that of the two-point function, which would scale as $\alpha^{-(D+n)}$ for the same power spectrum. This conclusion was already found from numerical observations in Obreschkow et al. (2013). Furthermore, compared to conventional correlation functions, the second relation is also somewhat unintuitive: increasing the number of objects (and hence the clustering) causes a decrease in the amplitude of $\ell(r)$ instead of an increase, as one would expect. The cause of this feature is that the increased number of different translations and rotations of objects randomise the phase factors, hence decreasing their correlation. This effect can also be understood by employing the halo model for a particular class of objects with orientation $\boldsymbol{\phi}$, mass m and density profile $\delta_H(\boldsymbol{r}\,|\,\phi, m)$. For random positions of the halos, one only needs to consider the one-halo term, so that, according to Smith and Watts (2005), the power-

[2] All objects are assumed to have random positions and orientations.

and bispectrum become

$$P(k) = (2\pi)^D \int \mathrm{d}m \, \frac{n(m)\, m^2}{\overline{\rho}^2} \int \frac{\mathrm{d}\boldsymbol{\phi}}{\Omega_D} \, |U_{\boldsymbol{k}}(\boldsymbol{\phi}, m)|^2 \,, \tag{4.9}$$

$$B(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) = (2\pi)^D \int \mathrm{d}m \, \frac{n(m)\, m^3}{\overline{\rho}^3} \int \frac{\mathrm{d}\boldsymbol{\phi}}{\Omega_D} \prod_{i=1}^{3} U_{\boldsymbol{k}_i}(\boldsymbol{\phi}, m) \,, \tag{4.10}$$

where $\overline{\rho}$ denotes the total mass density of all halos and Ω_D the surface area of the D-dimensional unit sphere. If all halos have identical mass M, the mass function is given by $n(m) = N/V\,\delta^D(m-M)$ and we see that $P \propto 1/N$, i.e. $B \propto 1/N^2$. Hence, the combination of power spectra and the bispectrum in Eq. (4.3) gives rise to the scaling $\sim 1/\sqrt{N}$.

Being a relatively new tool in large-scale structure statistics, the ILCF has been applied only a few times so far. For instance, in Obreschkow et al. (2013) it was shown to be a significantly more sensitive tool to analyse properties of dark matter than the two-point function. The authors examined a series of cosmological N-body simulations with different warm dark matter masses, finding that $\ell(r)$ distinguishes those masses out to scales approximately five times larger than the two-point function. This feature is a direct consequence of the ILCF's sensitivity to elongated structures and its ability to exclusively probe the non-linear stages of gravitational evolution. In another recent application (Alpaslan et al., 2014), the ILCF was used to identify a possibly new class of structure in the matter distribution, denoted as tendril galaxies. These galaxies were spotted in the *Galaxy and Mass Assembly* (*GAMA*, (Driver et al., 2009)) catalogue as remnants after all filament and isolated galaxies had been removed by appropriate detection algorithms. They appear in thin chains that extend into voids and display a different line correlation than those galaxies residing in the larger filaments. Finally, in Wolstenhulme et al. (2015), the ILCF was computed using tree-level perturbation theory giving good agreement with direct estimates from N-body simulations. An analytic formula for $\ell(r)$ allows it to be modelled properly and opens the door for measuring various cosmological parameters.

4.3 The anisotropic line correlation function

Our ability to measure cosmological parameters with the line correlation function is intimately linked to our understanding of how it is influenced by effects that cause anisotropies in the distribution of galaxies (see Sec. 4.4). Not knowing the systematics of these effects increases uncertainties in the result. However, an anisotropic signal does not only add complications, its detection would also open up new avenues to probe and constrain the cosmological evolution in our universe. For this reason, it is a vital

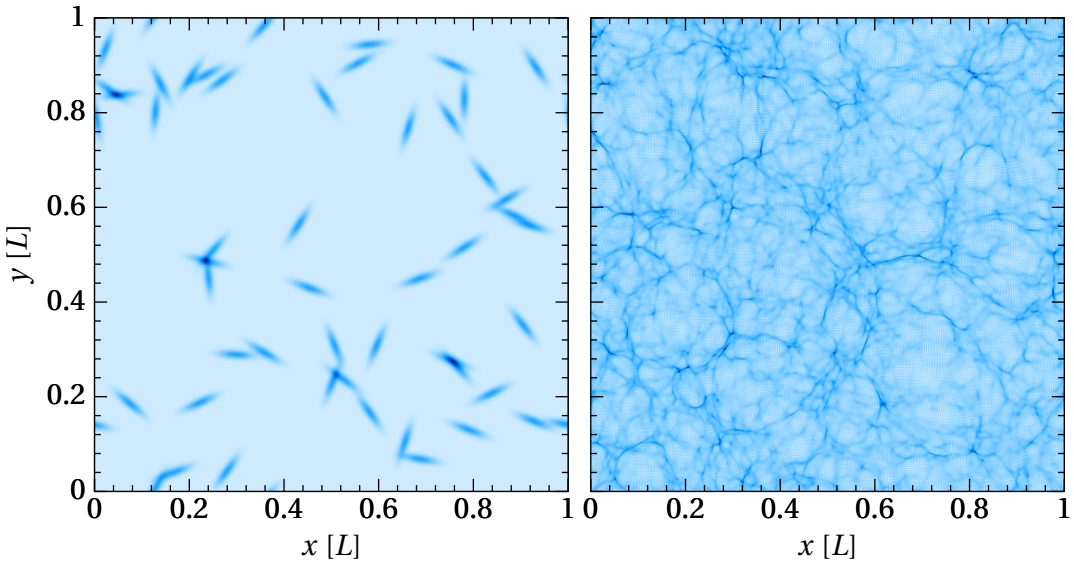


Fig. 4.1 Examples of two-dimensional density fields used throughout this chapter. Left: 50 aspherical Gaussian halos with $\sigma_x = 0.006\,L$, $\sigma_y = 0.024\,L$ and orientations drawn from a flat distribution. Right: Zel'dovich approximation, cosmological parameters are to be found in the text.

step to extend the ILCF to account for radial and transverse distance scales separately, allowing us to detect distortions along the line-of-sight.

Due to the corrections regularising the ILCF, there are various possibilities to break rotational symmetry. In the following we will therefore explore two different mode cut-offs, one being spherical as in the definition in Eq. (4.2), the other one aspherical. We show that the latter is better suited to pick up anisotropies and test some of its properties.

4.3.1 A spherical cut-off

Naively, it is straightforward to write down a two-dimensional analogue of the ILCF, especially when adopting the distant-observer approximation, so that we can identify the line-of-sight direction with the z-direction – we simply restrict the angular average in our prescription in Eq. (4.4) to orientations of the transverse part of the vector $\boldsymbol{r}$, giving

$$\ell(r_\perp, r_\parallel) = \frac{V^2}{(2\pi)^{2D}} \left(\frac{r^D}{V}\right)^{3/2} \iint\limits_{\substack{|\boldsymbol{k}_1|,\,|\boldsymbol{k}_2|,\\ |\boldsymbol{k}_1+\boldsymbol{k}_2| \le 2\pi/r}} \mathrm{d}^D k_1\, \mathrm{d}^D k_2\, \frac{\delta_{\boldsymbol{k}_1}\,\delta_{\boldsymbol{k}_2}\,\delta_{-\boldsymbol{k}_1-\boldsymbol{k}_2}}{|\delta_{\boldsymbol{k}_1}\,\delta_{\boldsymbol{k}_2}\,\delta_{-\boldsymbol{k}_1-\boldsymbol{k}_2}|} \times \cos\left[(k_{\parallel,1} - k_{\parallel,2})\, r_\parallel\right] w_{D-1}\left(|\boldsymbol{k}_{\perp,1} - \boldsymbol{k}_{\perp,2}|\, r_\perp\right). \tag{4.11}$$

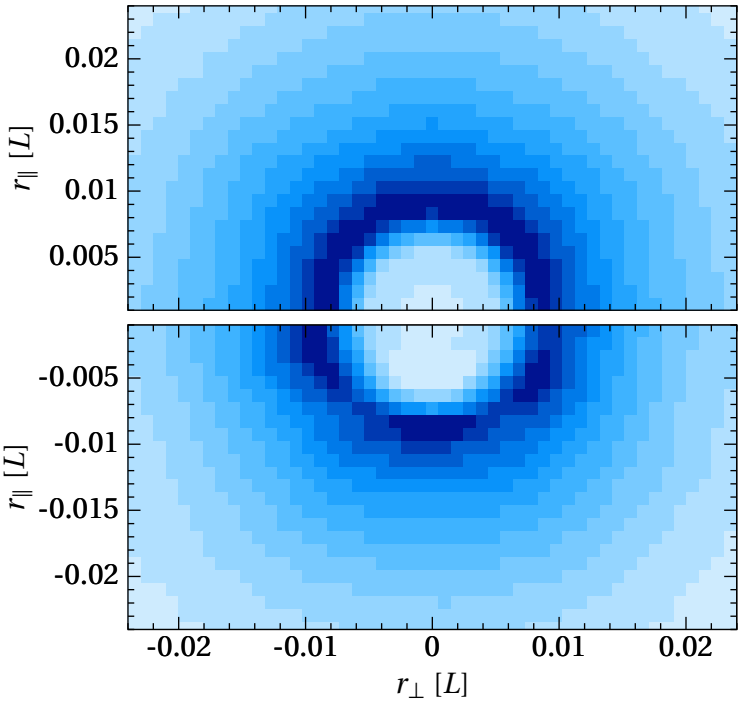


Fig. 4.2 Anisotropic line correlation function with spherical cut-off for statistically isotropic (top) and anisotropic fields with von Mises parameters $\kappa = 3$, $\mu = 0$ (bottom). Each field contains 50 halos and the width of their Gaussian profiles are chosen as $\sigma_x = 0.006\,L$ and $\sigma_y = 0.024\,L$. Results are averaged over 50 different realisations. Colour scales are adjusted.

Here, $r_\perp \equiv \sqrt{x^2 + y^2}$ and $r_\parallel \equiv z$ ($D = 3$) denote transverse and radial separations, respectively.

To see how well the function above distinguishes between statistically isotropic and anisotropic density fields, we consider a simple test case. We construct fields consisting of superpositions of aspherical halos with a Gaussian density profile (see upper panel in Fig. 4.1). While their centre positions are randomly distributed, their orientations are either drawn from a flat or a von Mises probability distribution. The von Mises distribution (Mardia and Jupp, 1999; Watts et al., 2003) is a close approximation of the wrapped normal distribution — the analogue of the usual normal distribution on a circle — and depends on two parameters: μ, the centre of the distribution; and κ, which is a measure for its 'peakedness', i.e. increasing κ causes a higher concentration about the centre. It reads

$$p(\phi\,|\,\mu,\kappa) = \frac{\mathrm{e}^{\kappa\cos(\phi-\mu)}}{2\pi\, I_0(\kappa)}\,, \tag{4.12}$$

where $I_0(\kappa)$ denotes the modified Bessel function of zeroth order. Furthermore, all fields here and in the remainder of this chapter, if not stated otherwise, are set up in a

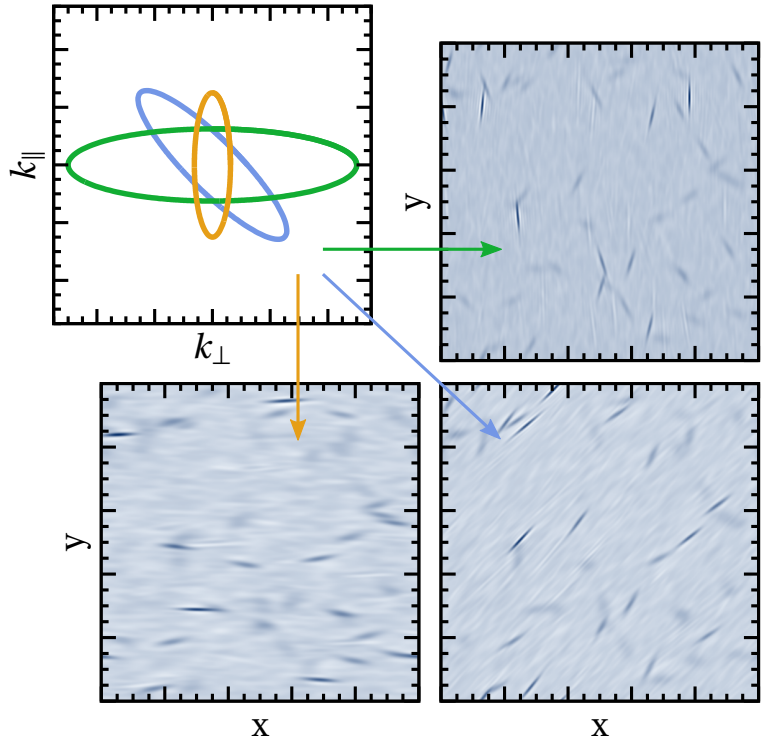


Fig. 4.3 Cut-off ellipses for different pairs $(r_\perp, r_\parallel)$ and the resulting whitened and filtered density fields. Note that the blurring increases from top right to bottom left because of a decreasing number of modes.

two-dimensional box of side length L with $N = 512$ grid cells (details are given in the captions of the corresponding figures). Extracting the line correlations of these sample fields according to Eq. (4.11) leads to the plot in Fig. 4.2. While $\ell(r_\perp, r_\parallel)$ only depends on $r = \sqrt{r_\parallel^2 + r_\perp^2}$ for the statistically isotropic fields, as expected, we do not observe a clear deviation from rotational symmetry in case of the anisotropic ones. This implies that the anisotropic signal is smeared over the whole $(r_\perp, r_\parallel)$-plane, which renders the spherical cut-off either tricky or ill-suited for the purposes given in the introduction of this section. For that reason we consider an alternative method that incorporates an aspherical cut-off.

4.3.2 An aspherical cut-off

When exploiting our freedom to alter the mode cut-off, we must still ensure that it regularises the integral for large scales and at the same time does not introduce any artificial anisotropies. Therefore, we impose that it should satisfy the following two conditions:

1. The number of enclosed modes scales with r^{-D};
2. The number of enclosed modes for each pair of values $(r_\perp, r_\parallel)$ with the same r is constant.

There are still many possible cut-offs that satisfy these conditions; the one that we propose here is based on oblate spheroids of constant eccentricity (i.e. ellipses in the two-dimensional $(k_\perp, k_\parallel)$-plane). The semi-minor axes is supposed to scale as $2\pi/r$ and more importantly, its orientation is chosen to be aligned with the vector $\boldsymbol{r} = (\boldsymbol{r}_\perp, r_\parallel)$, such that the spheroid rotates for varying scales. This choice is motivated by the fact that by filtering out all modes except those that predominantly belong to a given direction in k-space, objects are singled out that are aligned with the transverse direction in real space. This is demonstrated for a two-dimensional example in Fig. 4.3 which shows three different cut-offs and the corresponding whitened and filtered density fields. The green ellipse, for instance, belongs to a configuration where $r_\perp = 0$.

Mathematically, this rotating cut-off may be expressed as

$$\theta_\eta(\boldsymbol{k}, \boldsymbol{r}) \equiv k^2 r^2 + \left(\eta^2 - 1\right)(\boldsymbol{k}\cdot\boldsymbol{r})^2 \leq 4\pi^2\,, \tag{4.13}$$

where $\eta = \text{const.} \geq 1$. In two dimensions, we can convert Eq. (4.13) into polar form with φ the polar angle in the $(k_\perp, k_\parallel)$-plane and we get

$$k \leq \frac{2\pi/(\eta r)}{\sqrt{1 - \left(1 - \frac{1}{\eta^2}\right)\left(\frac{r_\parallel}{r}\cos\varphi - \frac{r_\perp}{r}\sin\varphi\right)^2}}\,. \tag{4.14}$$

Thus, both semi-axes scale with r^{-1} while the eccentricity of the ellipses is constant and given by $\varepsilon = \sqrt{1-\eta^{-2}}$, as desired. For $\eta = 1$ we recover the original spherical truncation, while different values of η influence the signal-to-noise ratio and the sensitivity to anisotropies, as we will see in the following.

Since the eccentricity is left as a free parameter, we may ask whether there is a particular choice of η that is optimal in the sense that it gives optimal signal-to-noise ratios for a range of scales. To determine if there is such a choice we employ the same sample of statistically isotropic fields as above, consisting of Gaussian halos with random positions and orientations. After computing the average line correlation and its variance from a number of realisations, we integrate out the angular dependency by averaging over all data points belonging to one of 25 radial bins of width $\Delta r \sim 0.0016\,L$ that extend outwards from the origin $r = 0$. Normalising the variance by the square of the average, we obtain an r-dependent noise-to-signal measure that we can compare for different values of η. The results of this analysis are shown in the upper panel of Fig. 4.4. All curves display a nearly continuous increase in the noise level, which is a consequence of a decreasing number of modes. However, the line correlation with spherical cut-off exhibits the largest noise-to-signal for most scales, indicating that a different choice for η is preferable. Indeed, we observe that $\sigma_\ell^2/\overline{\ell}^2$ drops when the cut-off becomes more elliptical and eventually reaches a minimum that roughly lies in the interval $\eta^2 \sim [16, 32]$.

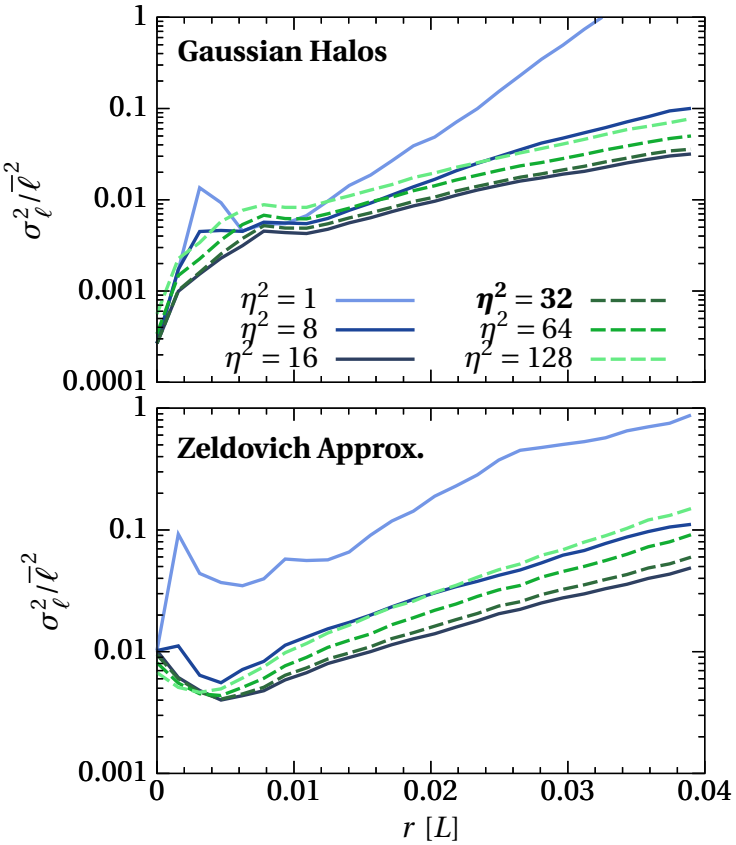


Fig. 4.4 Ratio of variance to squared average as a function of radial scale r based on 20 realisations. *Upper panel:* density fields consisting of Gaussian halos with random positions and orientations. *Lower panel:* fields set up according to the Zel'dovich approximation, also based on 20 realisations.

For larger values, the noise level begins to rise again, such that the optimal η must reside in the given interval. A possible explanation for this behaviour could be that the elliptic cut-off particularly enhances filamentary structures (see Fig. 4.3), which leads to an increase in signal as the configuration dependence of the LCF preferentially targets such structures. For a certain value of η this enhancement must saturate because beyond that value the cut-off ellipses become so thin that they tend to blur out the resulting filtered density field, which weakens the signal. Since the geometry of the cut-off matches that of the Gaussian halos, one might suspect that the optimal choice of η is tied to this particular class of density fields. To check this suspicion, we employ a second, more realistic set of density fields which are set up according to the Zel'dovich approximation (Zel'dovich, 2009) for a ΛCDM universe with parameters $\Omega_m = 0.314$, $h = 0.674$, $\sigma_8 = 0.9$ and baryon fraction $f_b = 0.038$ in a box of side length $L = 1\,h^{-1}\,\mathrm{Gpc}$ (for an example see lower panel of Fig. 4.1). Applying the same analysis to the Gaussian halo fields leads to the results in the lower panel of Fig. 4.4. As before, we note that a minimal noise-to-signal ratio is obtained in the parameter range $\sim [16, 32]$, strengthening the

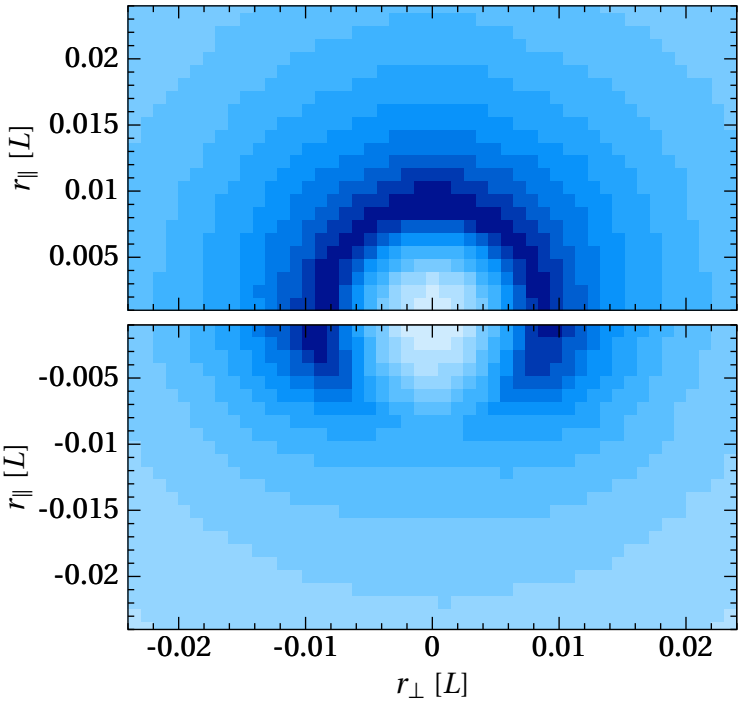


Fig. 4.5 Same as Fig. 4.2 but using the anisotropic line correlation function with aspherical, rotating cut-off. Colour scales are adjusted.

previous result[3]. Therefore, we define the anisotropic line correlation function (ALCF) for the remainder of this chapter by adopting the rotating cut-off with $\eta^2 = 32$. After fixing η, we still have the freedom to adjust the profile of the cut-off, for instance to a Gaussian instead of a top-hat profile, and thus minimise the noise-to-signal ratio even further. However, for simplicity, we work with the top-hat filter in this article.

Having settled on a definition of the ALCF, let us see how this choice influences the sensitivity to anisotropies. To that end, we repeat the test case from above where the fields are either statistically isotropic or anisotropic by giving the orientation of the Gaussian halos a strong preference in the line-of-sight direction. The resulting ALCFs are displayed in Fig. 4.5. While the isotropic fields still give rise to rotationally invariant line correlations, we now see a clear anisotropic signal in the lower panel. Since the isocontour lines appear squashed and the signal peaks close to $r_\parallel = 0$, the whole transverse direction is enhanced. Given that most of the halos point in the line-of-sight direction, this feature may appear counter-intuitive, but it is simply a continuation of the properties of the ILCF mentioned earlier. In Sec. 4.2.2 it was found that the amplitude of the ILCF decreases when the number of objects increases, which is reproduced here as well. In the following section we study the properties of the ALCF more closely.

[3]This result was obtained in a $2D$ analysis; we caution the reader that the optimal value for η in $3D$ might be different.

4.3.3 Properties and examples

To characterise the properties of the ALCF and verify that it extends those of the ILCF in a consistent manner, we work again with the simple Gaussian halo fields and consider two anisotropic setups. For the first, we introduce a preferred direction, as already seen in the examples of the last two sections; for the second, we vary the scale of the halos depending on their orientation.

Varying number densities

As before, the preferred direction is established by drawing the orientation angles of halos from a von Mises distribution. We want to analyse how the strength of anisotropy, parametrised by the von Mises parameter κ, affects $\ell(r_\perp, r_\parallel)$ and in particular, whether there is a scaling relation similar to the one valid for the ILCF. Therefore, we firstly introduce a probability density of objects, $p_H(\phi)$, which is defined on the interval $[-\pi/2, \pi/2]$, such that

$$N \times p_H(\phi)\,\mathrm{d}\phi \equiv \text{number of objects in the bin } \mathrm{d}\phi\,, \tag{4.15}$$

$$\Rightarrow \quad \int_{-\pi/2}^{\pi/2} p_H(\phi)\,\mathrm{d}\phi = 1 \tag{4.16}$$

with N, the total number of objects. Due to the symmetric geometry of the halos, p_H is related to the probability density of orientation angles by

$$p_H(\phi) = p(\phi) + p(\phi+\pi) = \frac{\cosh\left[\kappa\,\cos\left(\phi-\mu\right)\right]}{\pi\, I_0(\kappa)}\,, \tag{4.17}$$

where we have substituted in the von Mises distribution from Eq. (4.12).

As already suggested by Fig. 4.5 and in agreement with the interpretation of line correlations, we expect that the ALCF scales inversely with the probability density of objects. To verify this assertion, we compute the ALCF for five ensembles, each with $\mu = 0$ and a different value for κ. Afterwards, we extract the dependence on the angle ϕ from the two-dimensional function $\ell(r_\perp, r_\parallel)$ by averaging over a fixed number of points with constant ϕ, but constrained to the radial interval $0 \le r \le R = 0.02\,L$ to exclude areas of large sample variance, i.e.

$$\ell(\phi) \equiv \frac{1}{R}\int_0^R \mathrm{d}r\,\ell(r\sin\phi, r\cos\phi)\,. \tag{4.18}$$

The results are displayed in Fig. 4.6; error bars originate from averaging over 100 realisations. We then fit this data to determine the scaling behaviour and to see whether we can recover the respective values of κ that were used as input parameters for the

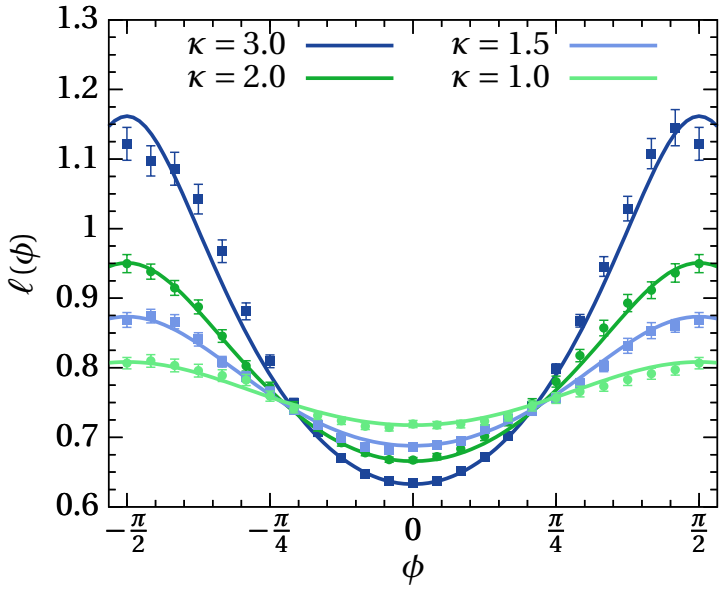


Fig. 4.6 Angular dependence of the ALCF for von Mises distributed density fields. Data points stem from the extraction procedure described in the text with error bars representing the 1σ uncertainty from averaging over 100 realisations. Solid lines are the corresponding fitting results.

Table 4.1 Input and fit parameters for von Mises distributed density fields. Parameters α and β are fitted simultaneously to all data sets, giving $\alpha = 0.58 \pm 0.02$ and $\beta = -0.23 \pm 0.02$.

	κ_1	κ_2	κ_3	κ_4	κ_5
Input	0.5	1.0	1.5	2.0	3.0
Fit	0.5 ± 0.1	1.1 ± 0.1	1.7 ± 0.1	2.3 ± 0.2	3.4 ± 0.2

construction of the fields with the model

$$\ell(\phi\,|\,\kappa, \alpha, \beta) = \alpha \left[\, p_H(\phi\,|\,\kappa) \,\right]^{\beta}, \tag{4.19}$$

where we allow for a varying amplitude and exponent. Since α and β should be universal (independent of κ), we perform a simultaneous χ^2-fit to all five data sets using a different fitting parameter κ_i and treating all data points as independent. The outcome of this fit along with the input parameters is summarised in Table 4.1, and shown for four example data sets as the thick solid lines in Fig. 4.6. We note that the fitted curves reproduce well the behaviour of the extracted data and we recover all peakedness parameters nearly within the error bounds. Moreover, taking the fitted values of α and β to make a prediction for fields with a flat probability density (i.e. $p_H = 1/\pi$), we compute $\ell = 0.76 \pm 0.03$ while getting $\ell = 0.753 \pm 0.001$ numerically. Hence, we can confirm that the ALCF scales with the probability distribution of objects where it is likely that the parameter β depends on the eccentricity of the cut-off ellipses.

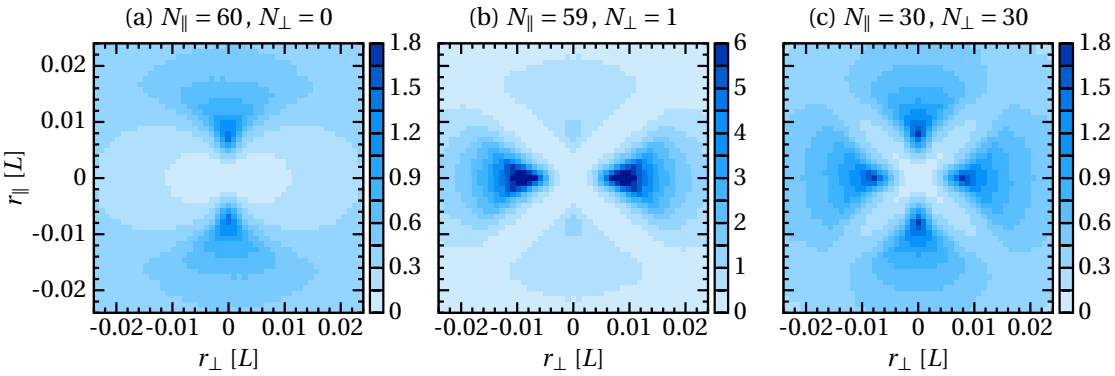


Fig. 4.7 Line correlations for fields with different numbers of elongated Gaussians pointing either along the line-of-sight ($N_\parallel$) or transverse direction ($N_\perp$), based on 20 realisations each. Note the change of colour scale in the middle panel.

Before moving on to the second test case we briefly describe another interesting feature related to the probability density of structures. To that end, we generate fields that can be considered as an extreme case of the von Mises fields above: varying numbers of elongated Gaussians pointing either in radial or transverse direction with fixed $N = N_\perp + N_\parallel$. From Fig. 4.7 we observe that a strong signal in the line-of-sight is received when all objects are pointing in that direction, see panel (a). Even so, this signal can be altered significantly if a single object is aligned with the transverse direction, see panel (b). Adjusting $N_\perp$ and $N_\parallel$ until equality leads to a symmetric signal, see panel (c). This feature of the ALCF can be explained along the same lines as above (see discussion below Eq. 4.8): in (a), no object is pointing in transverse direction, and the phases are consequently not correlated. In (b), however, phase correlations in transverse direction are generated by just a single object. This makes them dominant compared to the radial phase correlations, which are diluted due to the random superposition of multiple objects. Thus, care must be taken when using the plots of the ALCF to eyeball the strength of anisotropy.

Angularly varying scales

After having demonstrated the dependence on angular number densities, we turn our attention to angularly varying scales. Our sample halo fields are created as follows: for each orientation angle $\phi \in [-\pi/2, \pi/2]$, we determine the scale σ_y of the Gaussian profile from the function

$$\mathscr{S}(\phi\,|\,\overline{\sigma}, \Delta\sigma, n) = \overline{\sigma} + \Delta\sigma\,\cos\left(\omega\,\phi\right) \tag{4.20}$$

while keeping $\sigma_x = 0.006\,L$ fixed. As before, we intend to check how the ALCF is influenced by this anisotropy and especially how originally spherical lines of constant ℓ are deformed. In order to suppress any additional variations of the amplitude of $\ell(r_\perp, r_\parallel)$,

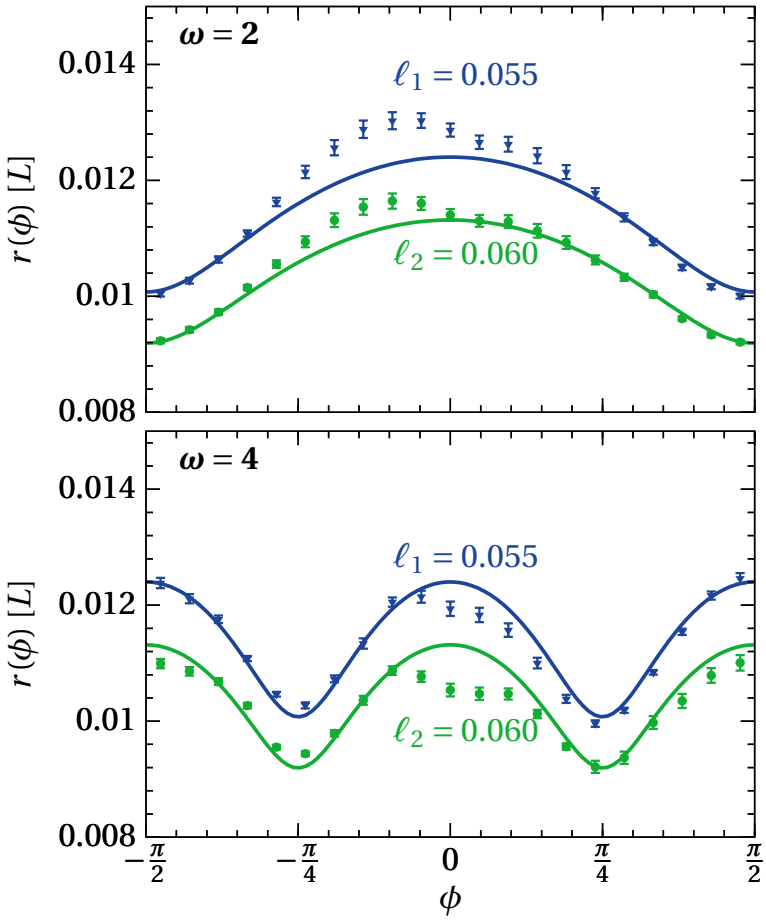


Fig. 4.8 Isocontours with 1σ error bars for two different values of ℓ extracted from the ALCF of fields with angularly varying scales, based on 100 realisations each. Solid lines stem from the results of the fitting procedure.

we need to hold the filling factor constant, so that

$$f(\phi) = p_H(\phi)\,\frac{V_0(\phi)}{L^D} \overset{D=2}{=} \frac{\pi\sigma_x}{L^2}\,p_H(\phi)\,\mathscr{S}(\phi) = \text{const.} \tag{4.21}$$

$$\Rightarrow \quad p_H(\phi) \propto 1/\mathscr{S}(\phi)\,, \tag{4.22}$$

where we resorted to the definition of f in Sec. 4.2.2 and used $r_0 = \sqrt{\sigma_x\sigma_y}$ for $D = 2$. Consequently, when determining the scales from $\mathscr{S}$, the respective orientation angles must be drawn from the distribution $1/\mathscr{S}(\phi)$. This distribution is realised using the method of inverse transform sampling (Devroye, 1986). With that, we set up two ensembles of fields for the cases $\omega = 2$ and $\omega = 4$, where we chose $\overline{\sigma} = 0.02\,L$ and $\Delta\sigma = 0.01\,L$. After evaluating the line correlation function we extract two isocontours for each ω, convert the Cartesian coordinates into 21 angular bins and average over 100 realisations which gives the data points in Fig. 4.8.

Table 4.2 Expected and fitted parameters for density fields with scales of explicit angular dependence.

	R_{l_1} $[10^{-2}\,L]$	R_{l_2} $[10^{-2}\,L]$	Δ	γ
Input	1.14 ± 0.02	1.03 ± 0.02	0.5	–
Fit	1.16 ± 0.04	1.06 ± 0.04	0.65 ± 0.06	0.13 ± 0.02

Since the average characteristic scale $\overline{r}_0 = \sqrt{\sigma_x \overline{\sigma}}$ is modified by the angle-dependent factor $\sqrt{1 + \Delta \cos(\omega\phi)}$ with $\Delta \equiv \Delta\sigma/\overline{\sigma}$, we may expect that any isocontour at scale R for a statistically isotropic field is distorted as

$$R \to R\left[1 + \Delta \cos(\omega\phi)\right]^{\gamma}, \tag{4.23}$$

with an unknown exponent γ which, as for the parameter β in the previous section, presumably depends on the eccentricity of the cut-off ellipses. We therefore fit the extracted contours with the model

$$r(\phi\,|\,R, \Delta, \gamma) = R\left[1 + \Delta \cos(\omega\phi)\right]^{\gamma}, \tag{4.24}$$

using different fitting parameters R_{l_i}, but equal Δ and γ throughout. The results from a simultaneous fitting procedure equivalent to the one above are represented by the solid lines in Fig. 4.8, with the corresponding best fit values given in Table 4.2.

We observe that the data is fitted well by the model, showing the expected distortion of radial scales, i.e. a stretching for those angles where the characteristic scale is larger. Furthermore, the fitted values for R match well to the same isocontours of statistically isotropic fields, which were estimated from an ensemble of randomly orientated halos with fixed scales $\sigma_x = 0.006\,L$ and $\sigma_y = \overline{\sigma} = 0.02\,L$. However, the parameter Δ is slightly off, which can be attributed to the scatter in the data and the fact that the fit is quite sensitive on both Δ and γ.

4.4 Application to redshift space distortions

Anisotropies in the distribution of galaxies occur naturally whenever the redshifts and positions (angles) acquired in a survey are converted into physical distances. They may be of geometric or kinematic nature, denoted by the AP effect and (dynamical) redshift space distortions, respectively. In this section, we analyse whether these observational effects are picked up by the ALCF and how these results compare to the ones from conventional two-point statistics. Since we primarily aim for a proof of concept, we will continue to use the two-dimensional halo and Zel'dovich mock fields introduced above.

4.4.1 The AP effect

The conversion of the redshift difference and angular separation of a pair of galaxies to a physical distance depends on two basic cosmological quantities: knowledge of the Hubble rate $H(z)$ is required for the computation of the line-of-sight separation; and the angular diameter distance $D_A(z)$ is required for the transverse separation,

$$r_\parallel = \frac{\mathrm{d}r}{\mathrm{d}z}\Delta z = \frac{\mathrm{d}}{\mathrm{d}z}\left(\int_0^z \frac{c\,\mathrm{d}z'}{H(z')}\right)\Delta z = \frac{c\,\Delta z}{H(z)}\,, \tag{4.25}$$

$$r_\perp = D_A(z)\,(1+z)\,\Delta\theta\,. \tag{4.26}$$

Consequently, creating a spatial map of the universe hinges on a cosmological model that has to be assumed a priori; if this model differs from the true cosmology, the deduced distances are wrong, causing distortions in the clustering signal. The first who proposed to use this effect as a probe for cosmological parameters were Alcock and Paczyński (Alcock and Paczynski, 1979). By considering an object whose intrinsic shape is known to be spherical, they deduced that measuring distortions from sphericity can be used to constrain the combination $H(z)\,D_A(z)$, and in turn, for instance, the equation of state parameter of dark energy. However, in practice, there are no objects, which are sufficiently spherical or whose length scales are known to sufficient accuracy, to perform the AP test on individually. Instead, one resorts to statistical standard rulers, that is, length scales that are statistically imprinted on large-scale structures, such as the one set by baryon acoustic oscillations (BAO); since BAOs are primarily a linear phenomenon, their signal is not considerably affected by non-linear physics. Thus, by measuring the position of the BAO peak and comparing it to the theoretical prediction, it is possible to conduct an AP test (Blake and Glazebrook, 2003). However, with the ever increasing volume coverage of galaxy surveys, an alternative has become feasible, in which a complete model of a given statistical measure, incorporating a possibly wrong choice of cosmology, is fitted to the data (see Ballinger et al., 1996; Okumura et al., 2008; Padmanabhan and White, 2008). A similar method could be employed based on the ALCF.

To quantify the impact of the AP effect on the line correlation function, we firstly parametrise the mismatch between the true and assumed cosmology by two squashing factors (following the notation of Ballinger et al., 1996):

$$f_\parallel = \frac{H_a(z)}{H(z)}\,, \qquad f_\perp = \frac{D_A(z)}{D_{A,a}(z)}\,, \tag{4.27}$$

where here and in the following the subscript 'a' indicates quantities derived from the assumed cosmology. From Eqs. (4.25) and (4.26) we see that assumed and true distance

scales are related by the matrix

$$\boldsymbol{r}_a = \mathscr{S}\,\boldsymbol{r}\,, \qquad \mathscr{S} = \begin{pmatrix} f_\perp^{-1} & & \\ & f_\perp^{-1} & \\ & & f_\parallel^{-1} \end{pmatrix}. \tag{4.28}$$

Due to number conservation, the over-density does not change in the assumed coordinate system, so that $\delta_a(\boldsymbol{x}_a) = \delta(\boldsymbol{x})$, which implies for the two-point function

$$\xi_a(\boldsymbol{r}_a) = \xi(\boldsymbol{r}) = \xi(\mathscr{S}^{-1}\,\boldsymbol{r}_a)\,. \tag{4.29}$$

Thus, any squashing of galaxy separations directly translates to an equivalent squashing of the isocontours of ξ. To compute the analogue effect for the ALCF we need to be more careful because of the scale dependent cut-off and prefactor. For that reason we transform $\ell_a(r_{\perp,a}, r_{\parallel,a})$ piece-wise and start from the whitened and filtered density field in the assumed coordinate system, which we write as a convolution between ϵ and the cut-off Θ in real space,

$$\begin{aligned} \epsilon_a * \Theta_{a,\boldsymbol{r}_a}(\boldsymbol{x}_a) &= V_a \int\limits_{\theta(\boldsymbol{k}_a,\boldsymbol{r}_a)\,\leq\,4\pi^2} \mathrm{d}^D k_a\, \mathrm{e}^{i\boldsymbol{k}_a\cdot\boldsymbol{x}_a} \epsilon_{a,\boldsymbol{k}_a} \\ &= V_a\,|\mathscr{S}|^{-1} \int\limits_{\theta(\mathscr{S}^{-1}\boldsymbol{k},\boldsymbol{r}_a)\,\leq\,4\pi^2} \mathrm{d}^D k\, \mathrm{e}^{i\boldsymbol{k}\cdot\mathscr{S}^{-1}\boldsymbol{x}_a}\, \epsilon_{\boldsymbol{k}}\,. \end{aligned} \tag{4.30}$$

Here, we used that the Fourier modes scale inversely to spatial scales and employed the identity $\epsilon_{a,\boldsymbol{k}_a} = \epsilon_{\boldsymbol{k}}$, which is another direct consequence of number conservation. Shifting the $\mathscr{S}$-matrix from $\boldsymbol{k}$ to $\boldsymbol{r}_a$ in the cut-off function causes an additional term that can be computed from the definition of θ in Eq. (4.13), yielding

$$\begin{aligned} \theta(\mathscr{S}^{-1}\boldsymbol{k}, \boldsymbol{r}_a) &= \theta(\boldsymbol{k}, \mathscr{S}^{-1}\boldsymbol{r}_a) + \left(f_\perp^2 - f_\parallel^2\right)\left(k_\perp^2\, r_{a,\parallel}^2 - k_\parallel^2\, r_{a,\perp}^2\right) \\ &= \theta(\boldsymbol{k}, \mathscr{S}^{-1}\boldsymbol{r}_a) + \delta_F(\boldsymbol{k}, \mathscr{S}^{-1}\boldsymbol{r}_a)\,, \end{aligned} \tag{4.31}$$

$$\delta_F(\boldsymbol{k}, \boldsymbol{r}) \equiv (1 - F^2)\left(F^{-2} k_\perp^2\, r_\parallel^2 - k_\parallel^2\, r_\perp^2\right), \tag{4.32}$$

where F is the ratio of $f_\parallel$ and $f_\perp$.[4] Plugging this result back into Eq. (4.30) gives

$$\begin{aligned} \epsilon_a * \Theta_{a,\boldsymbol{r}_a}(\boldsymbol{x}_a) &= V \int\limits_{\substack{\theta(\boldsymbol{k},\mathscr{S}^{-1}\boldsymbol{r}_a) \\ +\,\delta_F(\boldsymbol{k},\mathscr{S}^{-1}\boldsymbol{r}_a)\,\leq\,4\pi^2}} \mathrm{d}^D k\, \mathrm{e}^{i\boldsymbol{k}\cdot\mathscr{S}^{-1}\boldsymbol{x}_a}\, \epsilon_{\boldsymbol{k}} \\ &\equiv \epsilon * \widetilde{\Theta}_{F,\mathscr{S}^{-1}\boldsymbol{r}_a}(\mathscr{S}^{-1}\boldsymbol{x}_a)\,, \end{aligned} \tag{4.33}$$

[4] Note that δ_F arises purely as a consequence of the first term in Eq. 4.13. Eq. 4.31 and all of the following expressions are hence equally valid for the original spherical cut-off.

where we identified $V_a/|\mathscr{S}|$ with the true volume and let $\widetilde{\Theta}_F$ denote the altered cut-off that reverts to the original one for $F = 1$. Having thus determined the relation between the whitened and filtered density field in the assumed cosmology to the one in the true model, we still need to transform the prefactor. Using Eq. (4.28) again, we arrive at

$$\begin{aligned} r_{a,\parallel}^2 + r_{a,\perp}^2 &= f_\parallel^{-2}\, r_\parallel^2 + f_\perp^{-2}\, r_\perp^2 = \left(f_\parallel^{-2} + f_\perp^{-2}\right) r_\parallel^2 + f_\perp^{-2}\, r^2 \\ &= f_\perp^{-2}\, r^2 \left[1 + \left(F^{-2} - 1\right) \mu^2\right], \end{aligned} \tag{4.34}$$

where μ is the cosine of the angle to the line-of-sight, i.e. $\mu = r_\parallel / r$. All in all, we get

$$\ell_a(\boldsymbol{r}_a) = F^{\frac{3}{2}} \left[1 + (F^{-2} - 1)\left(\mathscr{S}^{-1}\mu_a\right)^2\right]^{\frac{3D}{4}} \widetilde{\ell}_F\left(\mathscr{S}^{-1}\boldsymbol{r}_a\right), \tag{4.35}$$

where

$$\left(\mathscr{S}^{-1}\mu_a\right)^2 \equiv \frac{f_\parallel^2\, r_{a,\parallel}^2}{f_\parallel^2\, r_{a,\parallel}^2 + f_\perp^2\, r_{a,\perp}^2}, \tag{4.36}$$

and the notation $\widetilde{\ell}_F$ refers to the line correlation function where all cut-offs have been modified by the function δ_F as in Eq. (4.33). Eq. (4.35), which is the main result of this section, relates the correlator for the assumed cosmology at scale $\boldsymbol{r}_a$ to the modified (but true) one at the shifted position $\mathscr{S}^{-1}\boldsymbol{r}_a$. As in Eq. (4.29) for the two-point function, it encodes the impact of the AP effect on the ALCF and can either be used as a basis for fitting existing data to extract $f_\perp$ and $f_\parallel$ or for simulating it. Since the effect is not immediately evident by inspecting Eq. (4.35), we compute the assumed line correlation function from a set of true fields with varying values for the squashing parameters.

To visualise the AP effect and enable comparison with the two-point function we decompose the data in the $(r_\perp, r_\parallel)$-plane into multipoles. As our density fields are two-dimensional, the n-th order multipole is calculated via

$$f_n(r) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(r\sin\phi, r\cos\phi)\cos\left(n\phi\right)\mathrm{d}\phi, \tag{4.37}$$

where f is either ξ or l and the sign convention is chosen such that an enhancement of the transverse direction compared to the line-of-sight translates into a positive signal for the quadrupole f_2. Using Eq. (4.35), we determine the monopole and quadrupole for an ensemble of Gaussian halo and Zel'dovich density fields with varying $f_\parallel$, but fixed $f_\perp$. The corresponding quadrupole-to-monopole ratios are plotted in the upper panels of Fig. 4.9. We notice that in both cases the AP effect introduces a clearly visible anisotropy by means of an increasing quadrupole signal with rising $f_\parallel$. Interestingly, the quadrupole is positive for small scales, meaning that the ALCF is squashed. However, with increasing scale the quadrupole changes sign, turning the initial squashing into a stretching of the

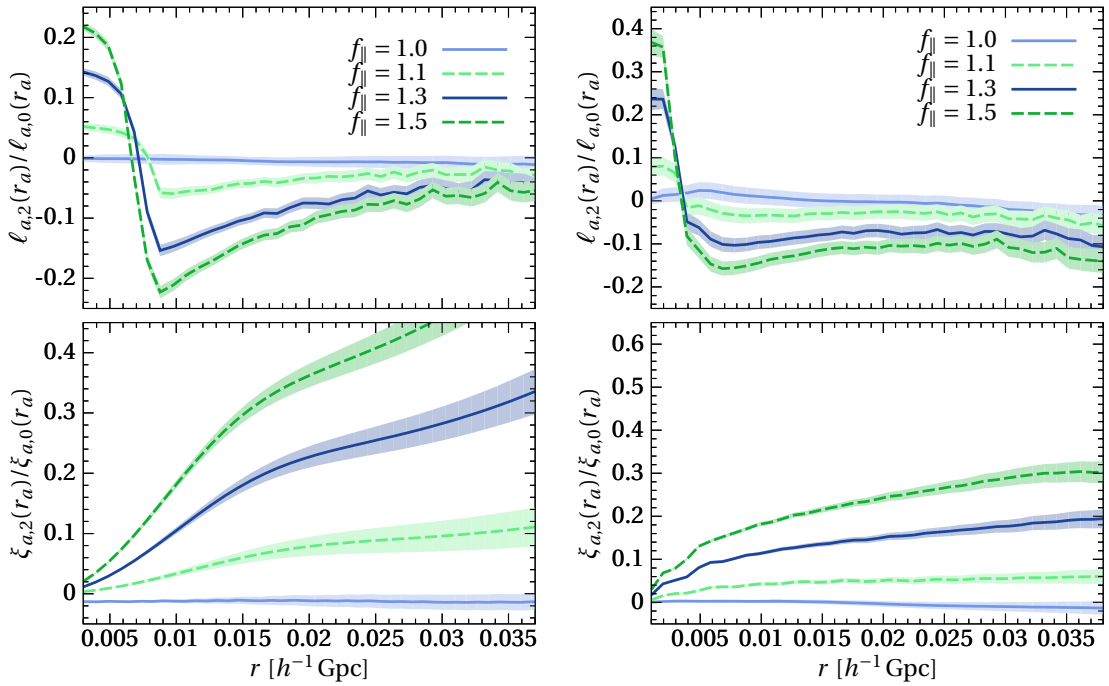


Fig. 4.9 *Left:* Quadrupole-to-monopole ratio for the ALCF (upper panel) and two-point function (lower panel), derived from a set of 50 Gaussian halo fields. The AP effect is simulated via Eq. (4.35) with $f_\parallel$ ranging between 1 and 1.5 while $f_\perp = 1$. Shaded areas mark 1σ uncertainties. *Right:* Same, but based on 50 realisations of Zel'dovich density fields.

line-of-sight. After attaining a maximum, the amplitude of the quadrupole decreases, but remains negative throughout. In comparison, the quadrupole-to-monopole ratio of the two-point function, which is displayed in the lower panels of Fig. 4.9, is positive for all scales, reflecting the squashing expected from Eq. (4.29). Although being unintuitive, the behaviour of the ALCF can be understood in light of the discussion in Sec. 4.3.3. If the separations of galaxies are squashed in a given direction, the scale and orientation of the structures in the density field change. In the case considered here (squashing in the line-of-sight direction) a preference for structures to be aligned with the transverse direction arises. This results in the enhanced radial signal since the ALCF scales inversely with the number density of objects.

While the ALCF is indeed capable of detecting the AP effect, the two-point function appears to be superior in telling apart different $f_\parallel$ from each other for both sets of fields. Its quadrupole-to-monopole ratio increases with scale, which renders ξ sensitive to $f_\parallel$ on all scales. On the other hand, the 1σ error regions begin to overlap at $r \sim 0.02\,L$ for the line correlation function, which implies that estimates based on ξ will be statistically more significant. This is a reasonable result owing to the number of modifications in Eq. (4.35), which tend to influence the ALCF in converse directions. Assuming $F > 1$ in the following, we first note that the modified cut-off gives rise to an increase in the ALCF's amplitude in the line-of-sight direction, as can be verified

empirically. Conversely, the distortion of the prefactor attains a minimum for $r_{a,\perp} = 0$ and accordingly strengthens the transverse direction compared to the line-of-sight. The same is true for the modification of scales by the matrix $\mathscr{S}^{-1}$, given that $\ell(\boldsymbol{r})$ is a decreasing function of $|\boldsymbol{r}|$. As a consequence, the net effect of the AP squashing on the ALCF is smaller than for the two-point function, where only one kind of distortion effect is present.

However, we would like to reiterate that the ALCF uses phase information, which is complementary to the amplitude information retained in the two-point correlator. Thus, both estimators can be used in conjunction in realistic applications to tighten constraints.

4.4.2 Kinematical redshift space distortions

Another consequence of measuring redshifts instead of physical distances are distortions due to the peculiar velocity of galaxies. Apart from the Hubble flow, they also contribute to the redshift and thus cause the galaxies to appear displaced along the line-of-sight.

Since galaxies result from the structure formation process through gravity, their peculiar velocities are induced by gravity as well. On large scales, where structures have not yet fully collapsed, all matter tends to fall into the nearest over-density. For that reason, a given galaxy that resides on the near-edge of a large cluster tends to move away from us, increasing its redshift. Consequently, its apparent position is closer to the centre of the over-density. A galaxy on the far-edge of the cluster would behave in the opposite way, so that it appears to be closer to us than it actually is. We therefore observe that structures on those large scales look squashed in the radial direction which is known as the Kaiser effect (Kaiser, 1987). He showed that in the distant observer limit, the power of scales that are aligned with the line-of-sight is boosted by the factor $1 + f\mu^2$ (μ being the cosine to the line of sight), so that the redshift-space dark matter power spectrum is given by

$$P_s(k) = (1 + f\,\mu^2)^2\, P(k)\,. \tag{4.38}$$

The amplitude of the distortion depends on the logarithmic growth rate f, which is related to the matter content in the universe and, according to Peebles (1980), well approximated by the power law $f \approx \Omega_m^{0.55}$. Thus, measuring the Kaiser effect enables us to put constraints on the matter density, or, if combined with an independent measurement of Ω_m, on the theory of general relativity.

At much smaller, nonlinear scales, objects within already collapsed structures acquire random virialised velocities which smear the structure in the line-of-sight direction leading to a characteristic shape that seems to point at the observer. Somewhat

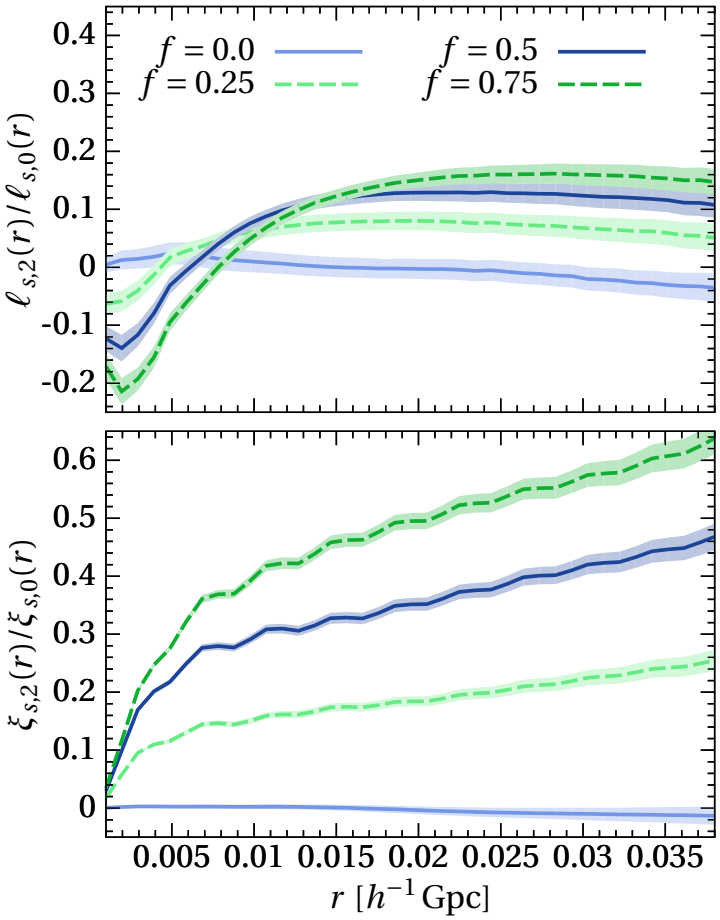


Fig. 4.10 Quadrupole-to-monopole ratio for the ALCF and two-point function calculated from a set of 50 Zel'dovich density fields incorporating the Kaiser effect. The amplitude of the distortion is parametrised by f.

misleadingly, these shapes were called *Fingers-of-God* (Jackson, 1972). Lacking an analytical description, several empirical models exist that try to mimic this small-scale effect by smearing the density field along the line-of-sight with the probability density function of velocities (Hawkins et al., 2003; Jackson, 1972). The models mainly differ in the assumption of how these velocities are distributed; commonly, it is expected that they acquire a Maxwellian distribution, which gives rise to a multiplicative damping term for the power spectrum of the form

$$F_{\mathrm{FoG}}(k,\mu) = \left(1 + \frac{k^2\mu^2\sigma_p^2}{2}\right)^{-1} . \tag{4.39}$$

The pairwise velocity dispersion σ_p is approximately of order $\sigma_p \sim 400\,\mathrm{km/s}$ (Hawkins et al., 2003).

The goal for the remainder of this section is to estimate how the ALCF is influenced by these kinematical redshift-space distortions. Hence, we take appropriate sets of

mock fields that simulate the Kaiser or Fingers-of-God effect and proceed as in the preceding section by computing the quadrupole-to-monopole ratio. In the Zel'dovich approximation the velocities are simply proportional (with proportionality factor f) to the displacement field and hence the Kaiser effect can be easily incorporated by displacing particle positions by this additional line-of-sight component. This leads to distorted density fields that seem to be derived from redshift-space.

In Fig. 4.10, we plot the quadrupole-to-monopole ratios for both the ALCF and the two-point function, estimated from these fields where we leave f as a free parameter which varies between 0.25 and 0.75. While the two-point function displays the anticipated behaviour, i.e. a squashing in the line-of-sight direction that becomes more prominent for increasing f, the ALCF exhibits a more complex quadrupole signal. On small scales, the quadrupole is negative, indicating a stretching in the line-of-sight, before it changes sign at a crossover scale that seems to shift towards larger r for higher f. Thereafter, it remains positive, though the squashing is less pronounced than for the two-point function. Since the Kaiser effect shrinks the radial dimension of structures in the density field, the enhancement of the transverse direction of the ALCF appears reasonable based on the discussion in Sec. 4.3.3.

Since the Zel'dovich approximation cannot account for the Fingers-of-God effect, we create a simplified setup which mimics these small-scale distortions. Firstly, we draw random positions that are taken to be the centre points of halos. We assume that these halos have a spherical Gaussian density profile and accordingly draw particle positions in each halo from a Gaussian distribution. Thereafter, we determine velocities from a Maxwellian distribution and displace all particles proportionally to their velocity along the line-of-sight. Computing the quadrupole-to-monopole ratio gives the plot in Fig. 4.11, showing a distinctive negative signal which confirms the stretching of structures in the line-of-sight. In this simplified case, the Fingers-of-God effect establishes a clear preference of the line-of-sight, with no structures being aligned with the transverse direction. However, as we have seen in Sec. 4.3.3, the signal of the ALCF may change drastically if this is the case, so it would be interesting to see how the Fingers-of-God effect appears in a realistic N-body simulation that also contains filamentary structures in all directions.

4.5 Conclusion

In this chapter, we presented an extension to the recently proposed line correlation measure (Obreschkow et al., 2013) that enables the identification of anisotropies in a given density field. The line correlation function is an estimator that relies purely on the phases of the density field and thus probes information not contained in ordinary two-point statistics.

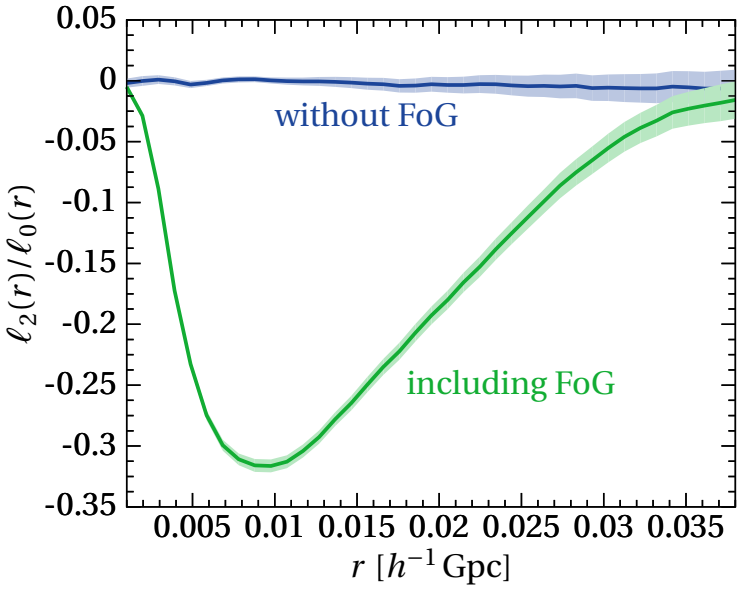


Fig. 4.11 ALCF quadrupole-to-monopole ratio for simple density fields that mimic the Fingers-of-God effect. Each field consists of 100 spherical halos, each of which contains 1000 particles that are Gaussian distributed about the centre with standard deviation $0.005\,L$. Their velocities are assigned via a Maxwellian distribution with velocity dispersion chosen such that the spread in position is $\sigma = 0.015\,L$. Results are based on 50 realisations.

Starting from a generalisation of the originally defined line correlation function, we noticed that anisotropies tend to be smeared over the plane of radial and transverse separations. For that reason, we introduced a novel mode cut-off scheme, which plays an integral part in the definition of the anisotropic line correlation function. We showed that this new function produces better signal-to-noise ratios and, by using simple test fields as replacements for real observational data, we demonstrated that it extends the properties of the original line correlation function in a consistent manner. In particular, we observed that the altered cut-off method allows for a clear distinction between statistically isotropic and anisotropic fields and we showed that it is sensitive to angularly varying number densities and scales of structures.

An interesting point that could be explored in future research is the optimisation of the cut-off function. While we have demonstrated that our choice of cut-off geometry with $\eta^2 = 32$ improves the signal-to-noise in the line correlation function for two morphologically different density fields, we have not shown that this filter actually maximises the signal-to-noise. To achieve that, one has to carry out a more systematic analysis involving the determination of the line correlation function covariance, which is beyond the scope of the work presented in this chapter.

In the second part, but still aiming primarily for a proof of concept, we considered observational effects, like the Alcock-Paczyński, Kaiser and the Fingers-of-God effect. We derived the relation between the assumed and true line correlation functions when

redshifts and angles of galaxies are transformed with a wrong prior cosmology and computed quadrupole-to-monopole ratios for all three effects, estimated from simple mock fields. We found that the line correlator is capable of detecting them, displaying different systematics than the two-point function. While its sensitivity appears to be somewhat worse than that of the two-point function, this is still a promising result: since the information probed by both measures is independent from each other, it is conceivable that employing both statistics can lead to tighter constraints on cosmological parameters than, for instance, a combination of the two- and three-point function.

To investigate these prospects, a more detailed study of redshift-space distortions is necessary. It is inevitable to employ N-body simulations to analyse both the Kaiser and Fingers-of-God effect in a realistic setup, and it would be interesting to see whether the crossover scale, which was observed in Fig. 4.10, is indeed a persistent feature. To understand this behaviour, additional analytical studies are desirable as well. A possible ansatz for such an investigation could be to reformulate the combination of bispectrum and power spectra that appears in the perturbative expression of the line correlation function (see Eq. (4.3)) in terms of the perturbative kernels in redshift-space which were derived in Scoccimarro et al. (1999a). Finally, it remains to be seen which parameters the line correlation function is able to constrain and to what accuracy.

In addition to the independence of information probed by the two-point and line correlation functions, both measures have different systematics, which allows the breaking of degeneracies in the parameter estimation. An interesting example is the logarithmic growth rate which describes the strength of the Kaiser effect (see Sec. 4.4.2). Due to the fact that we are measuring galaxy positions, which are biased tracers of the underlying dark matter distribution, the two-point function is actually only sensitive to the combination $\beta = f/b$, where b is the linear bias parameter. On the other hand, the line correlation function is independent of linear bias and hence constrains f directly, meaning that a joint analysis might allow for the determination of f and b separately. We leave these exciting new avenues for a future research project.

Appendices

Appendix 4.A Notes on the Implementation

In this appendix, we briefly comment on our numerical implementation of the line correlation function and show convergence tests.

We adopt a standard numerical discretisation scheme: the density field is contained in a finite cubic box of side-length L on which we impose periodic boundary conditions. The box is subdivided into an even number, N, of cells per side with the regular Cartesian grid spacing $\Delta x = L/N$. Accordingly, in Fourier space the box has a side-length $2\pi N/L$ with spacing $\Delta k = 2\pi/L$ and Nyquist frequency $k_{\mathrm{Ny}} = \pi N/L$. Upon discretisation using this scheme, integrals in configuration and Fourier space become

$$\int \mathrm{d}^D x\, f(\boldsymbol{x}) \quad \rightarrow \quad \left(\frac{L}{N}\right)^D \sum_{\boldsymbol{x}} f(\boldsymbol{x})\,, \tag{4.40}$$

$$\int \mathrm{d}^D k\, f_{\boldsymbol{k}} \quad \rightarrow \quad \left(\frac{2\pi}{L}\right)^D \sum_{\boldsymbol{k}} f_{\boldsymbol{k}}\,. \tag{4.41}$$

Applying these rules to the prescription of the line correlation function in Eq. (4.4) we get

$$\ell(r) = \left(\frac{r}{L}\right)^{\frac{3D}{2}} \sum_{\substack{|\boldsymbol{k}_1|,|\boldsymbol{k}_2|,\\ |\boldsymbol{k}_1+\boldsymbol{k}_2|\leq 2\pi/r}} w_D(|\boldsymbol{k}_1-\boldsymbol{k}_2|r)\, \frac{\delta_{\boldsymbol{k}_1}\delta_{\boldsymbol{k}_2}\delta_{-\boldsymbol{k}_1-\boldsymbol{k}_2}}{|\delta_{\boldsymbol{k}_1}\delta_{\boldsymbol{k}_2}\delta_{-\boldsymbol{k}_1-\boldsymbol{k}_2}|}\,. \tag{4.42}$$

Implementing this expression is straightforward, but computationally expensive due to the nested summations over $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$. Hence, we use the equivalent of Eq. (4.4) in configuration space, that is the spatially and rotationally averaged product of whitened density fields

$$\ell(r) = \left(\frac{r^D}{V}\right)^{\frac{3}{2}} \int \frac{\mathrm{d}^D x}{V}\, \langle \epsilon_\theta(\boldsymbol{x})\, \epsilon_\theta(\boldsymbol{x}+\mathscr{R}\boldsymbol{r})\, \epsilon_\theta(\boldsymbol{x}-\mathscr{R}\boldsymbol{r})\rangle_{\mathscr{R}}\,, \tag{4.43}$$

where $\mathscr{R}$ denotes a rotation matrix and the mode cut-off is written as ϵ_θ which stands for the convolution of $\epsilon(\boldsymbol{x})$ with an appropriate filter function $\Theta_{\boldsymbol{r}}$. Discretising the equation

above gives

$$\ell(r) = \left(\frac{r}{L}\right)^{\frac{3D}{2}} N^{-D} \sum_{\boldsymbol{x}} \langle \epsilon_\theta(\boldsymbol{x})\, \epsilon_\theta(\boldsymbol{x}+\mathscr{R}\boldsymbol{r})\, \epsilon_\theta(\boldsymbol{x}-\mathscr{R}\boldsymbol{r})\rangle_{\mathscr{R}} \tag{4.44}$$

$$= \left(\frac{r}{L}\right)^{\frac{3D}{2}} \mathrm{FFT}\left\{\langle \epsilon_\theta(\boldsymbol{x})\, \epsilon_\theta(\boldsymbol{x}+\mathscr{R}\boldsymbol{r})\, \epsilon_\theta(\boldsymbol{x}-\mathscr{R}\boldsymbol{r})\rangle_{\mathscr{R}}\right\}_{(\boldsymbol{k}=0)}, \tag{4.45}$$

where we replaced the summation over all $\boldsymbol{x}$ by the Fast-Fourier-Transform (FFT) algorithm. Since we can quickly transform back and forth between Fourier and configuration space using the FFT algorithm, it is advantageous to compute the convolution as well as the shift in position in Fourier space. There, shifting the field by the vector $\boldsymbol{r}$ amounts to multiplication with the phase factor $\exp(i\boldsymbol{k}\cdot\boldsymbol{r})$, while minimising rounding errors that might arise from the gridding when evaluating the field at a shifted position in configuration space. Hence, letting $\theta(\boldsymbol{k},\boldsymbol{r})$ either denote the original or modified cut-off in Eq. (4.13), we have

$$\epsilon_\theta(\boldsymbol{x}+\boldsymbol{r}) = \mathrm{FFT}^{-1}\left\{\epsilon_{\boldsymbol{k}}\, \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{r}}\, \theta(\boldsymbol{k},\boldsymbol{r})\right\}. \tag{4.46}$$

Eqs. (4.45) and (4.46) form the basis of all numerical computations in this article.

Let us consider a test case that allows us to compare with analytic results and study the convergence of the implementation with increasing N. We create density fields consisting of N_H spherical Gaussians with equal standard deviation σ. In this case (see Eqs. (4.9) and (4.10)),

$$\frac{B(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)}{\sqrt{P(|\boldsymbol{k}_1|)\,P(|\boldsymbol{k}_2|)\,P(|\boldsymbol{k}_3|)}} = \sqrt{\frac{V}{(2\pi)^D N_H}}, \tag{4.47}$$

such that for $D=2$ and, after integrating out one azimuthal angle, Eq. (4.3) gives

$$\ell(r) = 2\left(\frac{\sqrt{\pi}}{2}\right)^3 \left(\frac{r}{2\pi}\right)^3 \sqrt{\frac{V}{N_H}} \int_0^{2\pi/r} \mathrm{d}k_1\, k_1 \int_0^{2\pi/r} \mathrm{d}k_2\, k_2 \int_{-1}^{\mu_{\mathrm{cut}}} \frac{\mathrm{d}\mu}{\sqrt{1-\mu^2}}\, J_0(|\boldsymbol{k}_1-\boldsymbol{k}_2|\, r), \tag{4.48}$$

where $\mu = \cos\sphericalangle(\boldsymbol{k}_1,\boldsymbol{k}_2)$ and its upper limit is given by

$$\mu_{\mathrm{cut}} = \min\left\{1, \max\left\{-1, \frac{(2\pi/r)^2 - k_1^2 - k_2^2}{2k_1k_2}\right\}\right\}. \tag{4.49}$$

Eq. (4.48) can be integrated easily, for instance using a Monte-Carlo approach, to yield the analytic answer. However, to compare with the numerical result we need to take care of a subtlety which arises in the whitening process: whenever the amplitude of Fourier coefficients drops below a certain threshold value, their phase factors are not resolved properly any more; in order to avoid the introduction of artefacts, we need to cut out

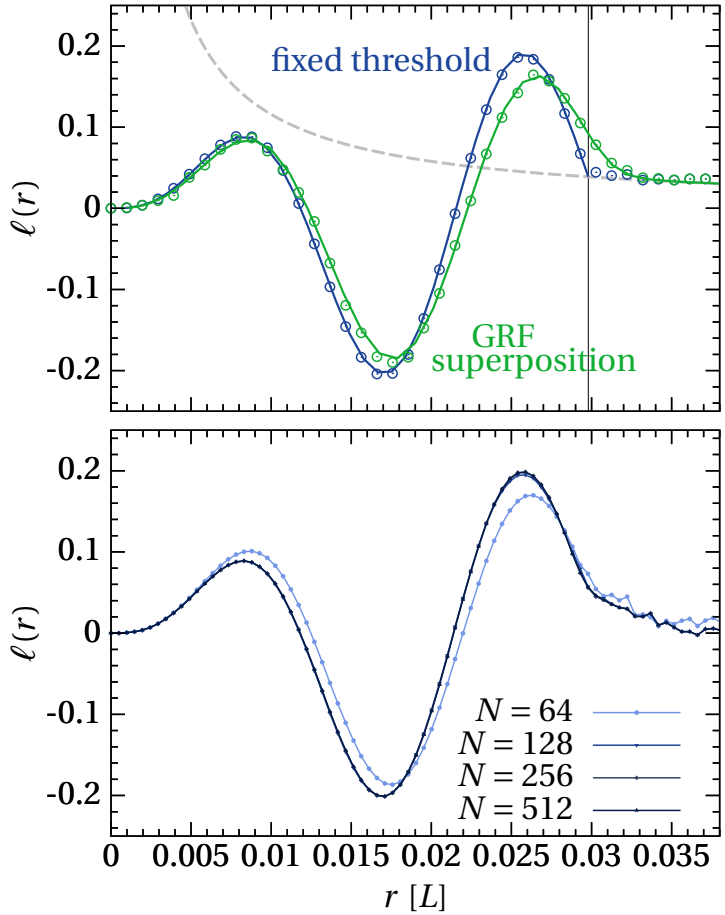


Fig. 4.A.1 *Top:* Comparison between analytics and numerics for a density field comprising $N_H = 100$ spherical Gaussian halos ($\sigma = 0.024\,L$). The dashed line is the result from Eq. (4.48) while the blue and green solid lines either account for a fixed threshold scale or the superposition with a Gaussian random field with $P_{GRF} = \text{const.}$; the corresponding data points stem from the purely numerical implementation averaged over 50 realisations. The vertical line indicates the threshold scale $r_{\text{th}} = 2\pi/k_{\text{th}}$. *Bottom:* Dependency on the number of grid points for a single realisation of the same density fields.

these coefficients[5]. We can apply the same rule as to the zero mode and set all modes with $|\delta_{\boldsymbol{k}}| < 10^{-7}$ to zero. For our test fields there exists an isotropic threshold scale k_{th} so that all modes $k > k_{\text{th}}$ have unresolved phase factors. Hence, this effect can be dealt with in Eq. (4.48) by replacing the integration limits for k_1 and k_2 by $\min\{k_{\text{th}}, 2\pi/r\}$ and similarly in μ_{cut}.

Another way to deal with unresolved phase factors is the superposition of the density field with a Gaussian random field (GRF) of small amplitude to ensures that Fourier coefficients with nearly vanishing amplitudes acquire random phases. This procedure guarantees that no artificial information compromises the signal: since the Gaussian

[5]This feature is due to the peculiar nature of the test fields considered here. Realistic density fields are unlikely to exhibit unresolved phase factors.

random field is uncorrelated to the Gaussian halos and assuming that it has zero mean, we only obtain a contribution to the overall power spectrum,

$$P(k) = P_G(k) + P_{GRF}(k)\,. \tag{4.50}$$

Doing so complicates Eqs. (4.47) and (4.48), but for known $P_{GRF}(k)$, it is still possible to find the analytic answer via Monte-Carlo integration.

The upper panel of Fig. 4.A.1 shows a comparison between numerical and analytic results. We see that to the right of the vertical line, which indicates the threshold scale $r_{\rm th} = 2\pi/k_{\rm th}$, the data points approach the dashed line, which is the analytic result based on Eq. (4.48). Below $r_{\rm th}$, the effect of cutting out unresolved phase factors becomes important and produces the oscillatory behaviour, which can be exactly reproduced for both methods if we account for the corresponding modifications in Eq. (4.48). Thus, this test case serves as a convincing consistency check between our numerical scheme and the perturbative expansion of the line correlation function.

Finally, the lower panel of Fig. 4.A.1 displays a convergence test for a varying number of grid points using a single realisation of the test fields above. The plot clearly shows that our method converges quickly, indicating only small discrepancies for N as low as $N = 128$. All of our computations were carried out with $N = 512$ grid points.

Contributions

The content of this chapter has been published in Eggemeier et al. (2015) and contains some work that was originally carried out in fulfilment of my Master's degree at the University of Göttingen. During my Ph.D. I significantly improved the numerical code used to estimate the line correlation function and revised and improved every step in the analysis. Moreover, I wrote the entire manuscript and prepared all figures for the paper, while my coauthors took on supervisory roles.

Chapter 5

Cosmology with phase statistics: parameter forecasts and detectability of BAO

This chapter extends the work presented in Chap. 4 by deriving predictions for the (isotropic) galaxy LCF, using a complete bias model at tree-level in perturbation theory and including shot noise. We demonstrate that the LCF is independent of linear bias at this order in perturbation theory, which breaks the degeneracy with the amplitude of fluctuations. Furthermore, we compute analytic expressions for its covariance matrix and confront all our predictions with measurements from a large ensemble of N-body simulations. Finally, we present Fisher forecasts for an idealised Stage III galaxy redshift survey of volume $V \sim 10\,h^{-3}\,\mathrm{Gpc}^3$ and out to $z = 1$, i.e. the same redshift coverage but only half the volume of *4MOST* (see Table 3.1). We find that, combining the LCF with the power spectrum and CMB priors from *Planck*, yields improvements up to a factor two for σ_8, b_1 and b_2, compared to using two-point information alone.

5.1 Introduction

The clustering of galaxies in our Universe is commonly analysed with one of the two simple statistical measures: the two-point correlation function or its Fourier space analogue, the power spectrum. Measurements of these statistics in the last decade led to the astonishing detection of BAOs in the clustering pattern of galaxies (Eisenstein et al., 2005; Tegmark et al., 2006) and subsequently to the increasingly precise constraints on our cosmological models which have confirmed the accelerated expansion rate of the present day Universe (see Alam et al., 2017, for a combination of two-point measurements from the final *BOSS* data release).

Over the next decades, further improvements are expected through developing theoretical advances that will enable us to extract information from deeper in the non-linear regime and through surveying larger volumes of space and so beating down the sample variance errors. However, the former is very challenging, especially as the scales involved approach shell crossing and the realm of uncertain baryonic physics effects takes hold, while the latter will come to a halt once surveys approach the cosmic variance limit. The path forward thus inevitably turns the question: what else should we measure to improve our understanding of the Universe? Higher-order statistics such as the three-point function, or equivalently the bispectrum, are obvious answers. For several decades the advantages of these measures have been known: the configuration dependence of the bispectrum allows us to break degeneracies present in the power spectrum, making it efficient in constraining galaxy bias (Fry, 1994; Matarrese et al., 1997); it was also shown to place improved constraints on other cosmological parameters (Sefusatti et al., 2006). Unfortunately, a wider application of these measures has been impeded by the slow development of improved theoretical modelling of these statistics and the challenge of accurate covariance matrices, which require very large sets of mock galaxy catalogues (see Gil-Marín et al., 2017; Slepian et al., 2017b, for recent measurements of three-point statistics, though). For that reason a number of simpler statistics have been proposed, which only measure a subset of the available three-point information and are designed to constrain certain parameters. These are for instance the bias estimators defined in Pollack et al. (2014) and Schmittfull et al. (2015) as well as the position dependent power spectrum (Chiang et al., 2014) and the related skew-spectra (Munshi and Heavens, 2010), which measure squeezed limits of higher-order statistics. Another idea that aims to compress bispectrum information into only a small number of modes was first explored in Regan et al. (2012).

In this chapter, we follow yet a different approach. As the power spectrum depends only on the squared amplitude of the Fourier mode, it is insensitive to the phase of the mode. Consequently, a measure purely based on these phases will strictly probe information that is not already contained in the power spectrum (Watts et al., 2003). Provided the density field is Gaussian, the phases remain random and a complete statistical description can be given in terms of the power spectrum. After initial perturbations start growing under the influence of gravity, though, correlations among the phases emerge, indicating a flow of information into higher-order moments of the density field. Phase information is therefore a direct probe of the non-linear regime and an appropriate measure for non-Gaussian information in the density field.

One phase based statistic that has emerged in recent times is the LCF (Obreschkow et al., 2013). It is defined as the three-point correlation function of the phases of the density field, whose three points are equally spaced on a line, each separated by a scale r. In Obreschkow et al. (2013) and in Chap. 4 it was shown that the LCF provides a

useful measure for quantifying the cosmic web, being able to differentiate between elongated, filamentary structures and node like structures (see also Alpaslan et al., 2014). Obreschkow et al. (2013) also showed that the LCF could be used to differentiate between cold and warm dark matter scenarios. Wolstenhulme et al. (2015) revealed the LCF's relation to conventional statistical quantities and showed that, at lowest order in standard perturbation theory (hereafter SPT), it can be expressed as a combination of bispectrum and power spectra, but that it in principle should contain information from cumulants of even higher order. Several further developments were made in Chap. 4, among which was the quantification of the effects of redshift space distortions on the LCF.

Even though the LCF is closely related to the bispectrum, it should be expected to exhibit different dependencies on cosmological parameters and also different covariance properties. Furthermore, it offers the possibility to measure in principal even higher-order information with just a three-point function. All that makes the LCF an interesting alternative to the standard methods, which is worth investigating more closely. This chapter thus aims to understand the response of the LCF to variations in the underlying cosmological models and to quantify the signal-to-noise of the estimates. Coupling together these quantities we aim to assess the cosmological parameter sensitivity of the LCF for upcoming galaxy redshift surveys that can be broadly classed as Stage III and Stage IV, in the language of the Dark Energy Task Force (Albrecht et al., 2006).

The chapter is broken down as follows: in Sec. 5.2 we provide a brief review of the LCF, we show how to compute the galaxy LCF and develop an analytic derivation of its covariance matrix. These theoretical predictions are then confronted with measurements from a large ensemble of N-body simulations in Sec. 5.4, before we move on to discuss the parameter sensitivity and error forecasts in Sec. 5.6, supported by a number of N-body simulations with varying cosmologies. Sec. 5.7 finally gives our conclusions.

5.2 Predictions from theory

5.2.1 The matter line correlation function

Consider a large volume of space within which is the realisation of a statistical homogeneous and isotropic random field. For a given scale r, the LCF of matter fluctuations can be defined as (for details see Obreschkow et al., 2013; Wolstenhulme et al., 2015):

$$\ell_{\rm m}(r) \equiv V^3 \left(\frac{r^3}{V}\right)^{3/2} \langle \epsilon_r(\boldsymbol{x})\epsilon_r(\boldsymbol{x}+\boldsymbol{r})\epsilon_r(\boldsymbol{x}-\boldsymbol{r})\rangle \tag{5.1}$$

where the first factor is a volume regularisation term, with V being the volume of the survey and $\epsilon_r(\boldsymbol{x})$ is the real space phase field smoothed on scale r. The Fourier transform

of the smoothed phase field can be written:

$$\epsilon_r(\boldsymbol{x}) = \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3}\epsilon_{\boldsymbol{k}}\mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{x}}W(k|r) \tag{5.2}$$

where the window function is a spherical top-hat in k-space: $W(k|r) = \Theta(1-kr/2\pi)$ and with $\Theta(x)$ being the Heaviside function. The phase field can now be readily defined as $\epsilon(\boldsymbol{k}) \equiv \delta(\boldsymbol{k})/|\delta(\boldsymbol{k})|$, with $\delta(\boldsymbol{k})$ being the Fourier transform of the over-density field. Note that the angle brackets in Eq. (5.1) indicate an averaging over an ensemble of random fields. Under the assumption of Ergodicity of the fields this becomes an average over volume and orientation of the direction vector of the line $\hat{\boldsymbol{r}}$ at each point in space.

On substitution of Eq. (5.2) into Eq. (5.1) we find that the line correlation can also be written:

$$\begin{aligned}\ell_{\mathrm{m}}(r) = \frac{V^3}{(2\pi)^9}\left(\frac{r^3}{V}\right)^{3/2} &\iiint\limits_{|\boldsymbol{k}_1|,|\boldsymbol{k}_2|,|\boldsymbol{k}_3|\leq 2\pi/r} \mathrm{d}^3k_1\,\mathrm{d}^3k_2\,\mathrm{d}^3k_3\ \langle\epsilon_{\boldsymbol{k}_1}\epsilon_{\boldsymbol{k}_2}\epsilon_{\boldsymbol{k}_3}\rangle \\ &\times \int \frac{\mathrm{d}\hat{\boldsymbol{r}}}{4\pi}\mathrm{e}^{i[\boldsymbol{k}_1\cdot\boldsymbol{x}+\boldsymbol{k}_2\cdot(\boldsymbol{x}+\boldsymbol{r})+\boldsymbol{k}_3\cdot(\boldsymbol{x}-\boldsymbol{r})]}\,,\end{aligned} \tag{5.3}$$

with the solid angle element $\mathrm{d}\hat{\boldsymbol{r}} \equiv \sin\vartheta\,\mathrm{d}\vartheta\,\mathrm{d}\varphi$. In order to proceed further one has to compute the ensemble average of the three phase factors,

$$\langle\epsilon_{\boldsymbol{k}_1}\epsilon_{\boldsymbol{k}_2}\epsilon_{\boldsymbol{k}_3}\rangle = \langle\exp\left[i\left(\theta_{\boldsymbol{k}_1}+\theta_{\boldsymbol{k}_2}+\theta_{\boldsymbol{k}_3}\right)\right]\rangle\ . \tag{5.4}$$

This expression can be evaluated by means of the joint probability density function (PDF) of Fourier phases, $\mathscr{P}(\{\theta_{\boldsymbol{k}}\})$, which was derived in Matsubara (2003, 2007). As is detailed in App. 5.A, for weakly non-Gaussian fields the PDF can be expanded in an Edgeworth series of higher-order correlators (in Fourier space poly-spectra) and using this result, one finds that to lowest order (Wolstenhulme et al., 2015):

$$\langle\epsilon_{\boldsymbol{k}_1}\epsilon_{\boldsymbol{k}_2}\epsilon_{\boldsymbol{k}_3}\rangle \approx \frac{(2\pi)^3}{V}\left(\frac{\sqrt{\pi}}{2}\right)^3 \frac{B(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)}{\sqrt{V\,P(\boldsymbol{k}_1)\,P(\boldsymbol{k}_2)\,P(\boldsymbol{k}_3)}}\,\delta_D(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)\,. \tag{5.5}$$

where the Dirac delta function δ_D appears in the above expression as a consequence of statistical homogeneity of the phase field, and where P and B are the power spectrum and bispectrum of the matter field, respectively. These are defined:

$$\langle\delta_{\boldsymbol{k}}\,\delta_{\boldsymbol{k}'}\rangle \equiv (2\pi)^3\delta_D(\boldsymbol{k}+\boldsymbol{k}')\,P(\boldsymbol{k})\,; \tag{5.6}$$

$$\langle\delta_{\boldsymbol{k}_1}\,\delta_{\boldsymbol{k}_2}\,\delta_{\boldsymbol{k}_3}\rangle \equiv (2\pi)^3\delta_D(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)\,B(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)\,. \tag{5.7}$$

The explicit dependence of Eq. (5.5) on the volume and hence the suppression of phase correlations in bigger surveys might appear surprising. However, with increasing volume

the density field will contain a greater number of halos and thus peaks in density. As has been shown in Hikage et al. (2004), the more peaks of comparable heights are enclosed in the sampling volume, the more the distribution of the phase sum will approach a uniform value. If this distribution becomes uniform, phase correlations are consequently diluted.

Finally on inserting Eq. (5.5) into Eq. (5.3) and integrating over $\boldsymbol{k}_1$ we find that at leading order, the line correlation can be written as an integral of the form:

$$\ell_m(r) \approx r^{9/2}\left(\frac{\sqrt{\pi}}{2}\right)^3 \iint\limits_{\substack{|\boldsymbol{k}_1|,\,|\boldsymbol{k}_2|,\\ |\boldsymbol{k}_1+\boldsymbol{k}_2|\,\leq\,2\pi/r}} \frac{\mathrm{d}^3k_1}{(2\pi)^3}\frac{\mathrm{d}^3k_2}{(2\pi)^3}\frac{B(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_1+\boldsymbol{k}_2)}{\sqrt{P(\boldsymbol{k}_1)\,P(\boldsymbol{k}_2)\,P(\boldsymbol{k}_1+\boldsymbol{k}_2)}}\int\frac{d\hat{\boldsymbol{r}}}{4\pi}\mathrm{e}^{i(\boldsymbol{k}_1-\boldsymbol{k}_2)\cdot\boldsymbol{r}}\,, \tag{5.8}$$

where after computing the integral we made the following relabellings $\boldsymbol{k}_2 \to \boldsymbol{k}_1$ and $\boldsymbol{k}_3 \to \boldsymbol{k}_2$. Finally, on computing the average over all orientations of $\boldsymbol{r}$ we arrive at the result:

$$\ell_{\mathrm{m}}(r) \simeq \left(\frac{r}{4\pi}\right)^{9/2} \iint\limits_{\substack{|\boldsymbol{k}_1|,\,|\boldsymbol{k}_2|,\\ |\boldsymbol{k}_1+\boldsymbol{k}_2|\,\leq\,2\pi/r}} \mathrm{d}^3k_1\,\mathrm{d}^3k_2\,\frac{B(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_1+\boldsymbol{k}_2)}{\sqrt{P(\boldsymbol{k}_1)\,P(\boldsymbol{k}_2)\,P(\boldsymbol{k}_1+\boldsymbol{k}_2)}}\,j_0\left(|\boldsymbol{k}_1-\boldsymbol{k}_2|\,r\right)\,, \tag{5.9}$$

where $j_0(x) = \sin x/x$ is the zeroth-order spherical Bessel function.

5.2.2 LCF from standard perturbation theory

On applying nonlinear SPT to the fluid equations for primordial Gaussian matter fluctuations in an expanding universe, one finds that at lowest order (tree-level) the bispectrum can be expressed in terms of the linear power spectrum P and a mode coupling kernel F_2 as (Bernardeau et al., 2002; Fry, 1984):

$$B(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) = 2\,F_2(\boldsymbol{k}_1,\boldsymbol{k}_2)\,P_1\,P_2 + \mathrm{cyc.}\,, \tag{5.10}$$

where in the above and in what follows we will use the short-hand notation $P_i \equiv P(k_i)$ and $P_{ij} \equiv P(|\boldsymbol{k}_i - \boldsymbol{k}_j|)$ and the power spectra are understood to be those obtained from linear theory. The mode coupling kernel F_2 reads:

$$F_2(\boldsymbol{k}_1,\boldsymbol{k}_2) = \frac{5}{7} + \frac{\hat{\boldsymbol{k}}_1\cdot\hat{\boldsymbol{k}}_2}{2}\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right) + \frac{2}{7}\left(\hat{\boldsymbol{k}}_1\cdot\hat{\boldsymbol{k}}_2\right)^2\,, \tag{5.11}$$

dropping its very weak dependence on the cosmological parameters (Bernardeau et al., 2002). On using again Eq. (5.5) and inserting Eq. (5.10) in this expression one can now choose to perform the integrations over the Dirac delta function so that the matter LCF

at tree level can be expressed as (Wolstenhulme et al., 2015):

$$\ell_{\rm m}(r) = 2\left(\frac{r}{4\pi}\right)^{9/2} \iint\limits_{\substack{|\boldsymbol{k}_1|,\,|\boldsymbol{k}_2|,\\ |\boldsymbol{k}_1+\boldsymbol{k}_2|\leq 2\pi/r}} \mathrm{d}^3k_1\,\mathrm{d}^3k_2\,F_2(\boldsymbol{k}_1,\boldsymbol{k}_2)\sqrt{\frac{P_1\,P_2}{P_{12}}} \times \Big[j_0\big(|\boldsymbol{k}_2-\boldsymbol{k}_1|\,r\big)+2\,j_0\big(|\boldsymbol{k}_1+2\boldsymbol{k}_2|\,r\big)\Big]. \tag{5.12}$$

While the above expression appears to require a 6-D integration, in fact it only requires a 3-D integration. This owes to the fact that all terms can be expressed in terms of the magnitudes of $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ and the angle between these vectors. Note that in deriving the above result we have assumed that for the denominator of Eq. (5.5) one-loop corrections to the power spectra are of negligible importance. As we will show in Sec. 5.4, for the range of scales that we will consider, this assumption appears to be reasonably accurate.

Finally, regarding the cosmological sensitivity of the matter LCF, we see from Eq. (5.12) that at tree-level the LCF is built from integrals over products of the linear matter power spectrum on different scales. Thus the cosmological dependence derives entirely from specifying the dependence of the linear power on cosmology. This can best be achieved with Boltzmann codes like `CAMB` (Lewis et al., 2000).

5.2.3 The galaxy line correlation function

We next turn to the issue of predicting the galaxy LCF (hereafter GLCF). There are two sources of non-Gaussianity that can contribute here, one stems from the relation between the galaxy over-density field δ_g and that of the matter – otherwise known as galaxy bias. The second from the fact that galaxies are a point sampled process and with the usual assumption that they share the same limiting properties for rare occupancy of micro-cells as a Poisson sampling process.

Firstly, for the galaxy bias, if we assume that the density field in real space obeys the cosmological principle, in that it is statistically homogeneous and isotropic on sufficiently large scales, one can argue that in this regime the relation between δ_g and δ should become linear (Fry and Gaztanaga, 1993; Smith et al., 2007). For this case the bias simply drops out of the equations for the GLCF, i.e. $\epsilon_g(\boldsymbol{x}) = \epsilon(\boldsymbol{x})$ and the galaxy and dark matter LCFs fully coincide.

On smaller scales non-linear terms enter the bias relation (Fry and Gaztanaga, 1993; Smith et al., 2007). Furthermore, if the relation does not depend on the present day local density but on the density of the initial patches, then non-local bias terms can contribute (Baldauf et al., 2012; Catelan et al., 2000; Chan et al., 2012; McDonald and Roy, 2009). For that reason we employ the Eulerian non-linear, non-local bias model that was proposed by McDonald and Roy (2009). In their model non-locality is introduced via a term which is quadratic in the tidal tensor $s_{ij}(\boldsymbol{x}) \equiv \left[\partial_i\partial_j\nabla^{-2} - \frac{1}{3}\delta^K_{ij}\right]\delta(\boldsymbol{x})$, such that

the galaxy over-density may be written as

$$\delta_g(\boldsymbol{x}) = b_1\,\delta(\boldsymbol{x}) + \frac{1}{2}b_2\Big[\delta^2(\boldsymbol{x}) - \langle\delta^2(\boldsymbol{x})\rangle\Big] + \frac{1}{2}b_{s^2}\Big[s^2(\boldsymbol{x}) - \langle s^2(\boldsymbol{x})\rangle\Big] + \ldots\,, \tag{5.13}$$

where $s^2 = s^{ij}s_{ji}$ and the dots indicate terms of even higher orders. The constants b_1, b_2, b_{s^2} denote the linear, non-linear and non-local bias terms, respectively, and the terms $\langle\delta^2\rangle$ as well as $\langle s^2\rangle$ ensure that $\langle\delta_g\rangle = 0$.

From the Fourier transform of Eq. (5.13) and by using PT we can readily determine the galaxy bispectrum. As before it can be expressed in terms of the linear dark matter power spectrum:

$$B_{\mathrm{g},123} = b_1^3\,P_1\,P_2\Big[2F_2(\boldsymbol{k}_1,\boldsymbol{k}_2) + c_2 + c_{s^2}\,S_2(\boldsymbol{k}_1,\boldsymbol{k}_2)\Big] + \mathrm{cyc.}\,, \tag{5.14}$$

where we used the short-hand notation $B_{\mathrm{g},123} \equiv B_\mathrm{g}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)$ and defined $c_2 \equiv b_2/b_1$ and $c_{s^2} \equiv b_{s^2}/b_1$, and where we have neglected all higher-order terms. We note that the non-local bias term has introduced an extra configuration dependence, which is encoded in the new kernel function (McDonald and Roy, 2009):

$$S_2(\boldsymbol{k}_1,\boldsymbol{k}_2) = \big(\hat{\boldsymbol{k}}_1\cdot\hat{\boldsymbol{k}}_2\big)^2 - \frac{1}{3}\,. \tag{5.15}$$

Instead of relating the galaxy phase field to the matter one via Eq. (5.13), we observe that the galaxy phase PDF can be obtained from the original PDF by replacing all spectra in its Edgeworth expansion with their biased equivalents. Hence, the three-point correlation of galaxy phase factors is the same as Eq. (5.5) with the dark matter quantities replaced by the galaxy bispectrum and power spectrum. As before we do not consider higher-order corrections to the denominator of Eq. (5.5), which are now also coming from non-linear and non-local bias, and simply substitute the linear galaxy power spectrum $P_{L,g} = b_1^2\,P_L$. We thus find that the GLCF can be written to lowest order as the sum of three terms:

$$\ell_\mathrm{g}(r) = \ell_\mathrm{m}(r) + c_2\,\ell_{b_2}(r) + c_{s^2}\,\ell_{b_{s^2}}(r)\,, \tag{5.16}$$

where each term can be written:

$$\ell_\alpha(r) = \Big(\frac{r}{4\pi}\Big)^{9/2} \iint\limits_{\substack{|\boldsymbol{k}_1|,\,|\boldsymbol{k}_2|,\\ |\boldsymbol{k}_1+\boldsymbol{k}_2|\,\leq\,2\pi/r}} \mathrm{d}^3k_1\,\mathrm{d}^3k_2\,\sqrt{\frac{P_1\,P_2}{P_{12}}}\Big[j_0\big(|\boldsymbol{k}_2-\boldsymbol{k}_1|\,r\big) + 2\,j_0\big(|\boldsymbol{k}_1+2\boldsymbol{k}_2|\,r\big)\Big]$$
$$\times\,\Gamma_\alpha(\boldsymbol{k}_1,\boldsymbol{k}_2)\,, \tag{5.17}$$

where the above equation holds for $\alpha \in \{\mathrm{m}, b_2, b_{s^2}\}$ and where

$$\Gamma_\alpha(\boldsymbol{k}_1, \boldsymbol{k}_2) = \left\{2F(\boldsymbol{k}_1, \boldsymbol{k}_2), 1, S_2(\boldsymbol{k}_1, \boldsymbol{k}_2)\right\}. \tag{5.18}$$

At this point we draw attention to one of the major advantages of the GLCF over more conventional galaxy clustering statistics, which is that in the limit of linear bias, i.e. where $b_2 = b_{s^2} = 0$, we see that $\ell_\mathrm{g}(r) = \ell_\mathrm{m}(r)$. Hence it is a direct probe of the matter distribution, independent of linear bias. We also notice that in the presence of non-trivial biasing the GLCF is sensitive only to the relative bias parameters, $c_2 = b_2/b_1$ and $c_{s^2} = b_{s^2}/b_1$.

Figure 5.1 (upper panel) shows all three terms. The dark matter LCF is indicated by the (black) thick line, and the dashed and dot-dashed lines represent the contribution of the local and non-local bias terms, respectively, where we set $c_2 = c_{s^2} = 0.5$. To evaluate the various expressions in Eq. (5.16) we adopted the same ΛCDM cosmology as described in Sec. 5.4 and generated the corresponding linear power spectrum using `CAMB`. We notice the different configuration dependence of all three functions, meaning that in principle it is possible to determine both c_2 and c_{s^2}, as well as the amplitude of fluctuations σ_8 from the LCF alone. However, the sole discernible difference between $\ell_\mathrm{m}(r)$ and $\ell_{c_{s^2}}(r)$ are the damped BAO wiggles in the latter, so we cannot expect the LCF to yield strong constraints on c_{s^2}. We stress that for the bispectrum there will be a remaining degeneracy between σ_8 and b_1 as can be seen from Eq. (5.14), which must either be broken via a joint analysis with the power spectrum or under the inclusion of a cosmic microwave background (hereafter CMB) measurement.

For the remainder of this work we will additionally assume that bias is local in Lagrangian space, in which case it can be shown that the non-local bias term is related to the linear one at first order (Baldauf et al., 2012; Chan et al., 2012),

$$b_{s^2} = -\frac{4}{7}(b_1 - 1)\,. \tag{5.19}$$

In this case galaxy biasing is only a function of two free parameters, b_1 and b_2, and as we will show these can be efficiently constrained through combination of the LCF with the power spectrum.

5.2.4 The effect of shot noise on the GLCF

Another additional source of non-Gaussianity that modulates the GLCF is sampling noise. Any real measurement from a galaxy survey will be compromised by shot noise – the fact that the matter field has to be reconstructed from a discrete and finite set of tracers that have been sampled from some underlying field which may be Gaussian. This effect increases as the density of objects in a given volume decreases and leads to

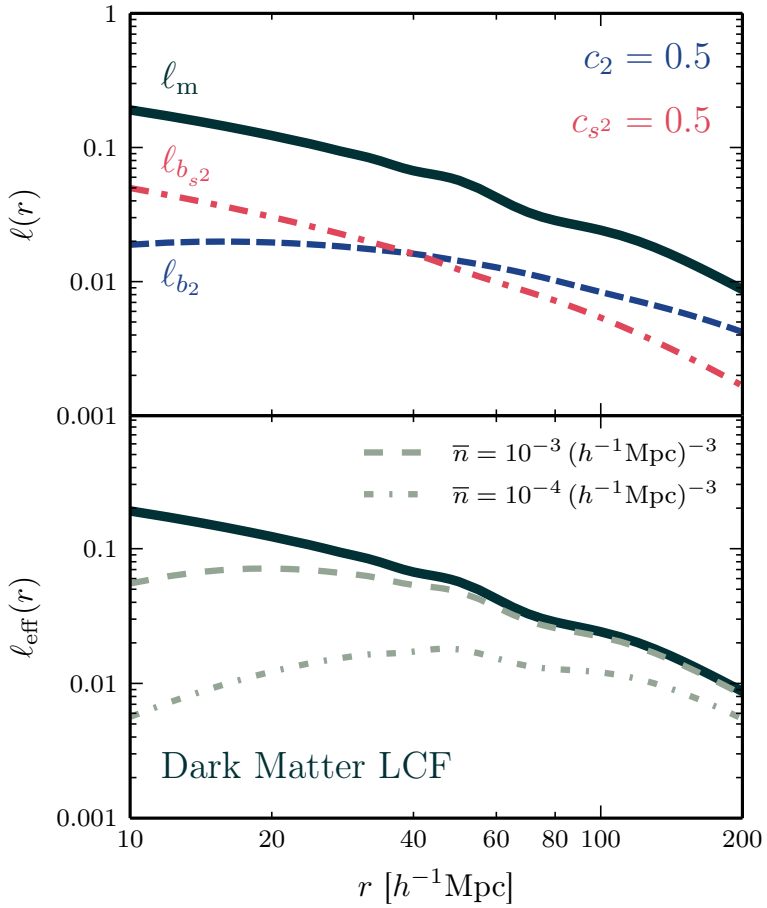


Fig. 5.1 *Upper panel:* Comparison of various contributions to the galaxy LCF with $\ell_{b_2}(r)$ in blue and $\ell_{b_{s^2}}(r)$ in red. The coefficients c_2 and c_{s^2} are both set to 0.5. *Lower panel:* Suppression effect from shot noise on the dark matter LCF, different line styles correspond to different number densities. The thick, black line in both panels is the LCF of the underlying matter field, i.e. for negligible shot noise.

an artificial enhancement in the clustering strength of galaxies in the power spectrum. In the absence of a selection function and finite survey geometry, a constant number density of tracers $\overline{n}$, would modulate the true power spectrum and bispectrum as follows (Matarrese et al., 1997; Peebles, 1980):

$$P^{\rm d}(k) = P(k) + \frac{1}{\overline{n}}\,, \tag{5.20}$$

$$B^{\rm d}(k_1,k_2,k_3) = B_{123} + \frac{1}{\overline{n}}\left[P_1+P_2+P_3\right] + \frac{1}{\overline{n}^2}\,, \tag{5.21}$$

where the superscript d stands for the discrete case. All shot noise terms involving factors of $1/\overline{n}$ can be subtracted to obtain an unbiased estimate of the true power spectrum or bispectrum, although they will still contribute to the errors.

To derive the effect of shot noise on the GLCF, we employ the same trick as employed in the last section. We assume the Fourier modes of the reconstructed matter field follow the same PDF as that of the true matter field, but with all spectra in the Edge-

worth expansion replaced by the corresponding discrete quantities. Accordingly, using Eqs. (5.20) and (5.21) the three-point phase correlator estimated from a set of discrete tracers is given by

$$\langle \epsilon_{\boldsymbol{k}_1}\epsilon_{\boldsymbol{k}_2}\epsilon_{\boldsymbol{k}_3}\rangle^{\rm d} = \frac{(2\pi)^3}{V}\left(\frac{\sqrt{\pi}}{2}\right)^3\sqrt{\frac{\nu_{\rm eff}(k_1)\,\nu_{\rm eff}(k_2)\,\nu_{\rm eff}(k_3)}{V\,P_1\,P_2\,P_3}}\left[B_{123}+\frac{1}{\overline{n}}\Big(P_1+P_2+P_3\Big)+\frac{1}{\overline{n}^2}\right] \times \delta_D(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)\,, \tag{5.22}$$

where we have defined

$$\nu_{\rm eff}(k) \equiv \frac{\overline{n}\,P(k)}{1+\overline{n}\,P(k)}\,. \tag{5.23}$$

This factor encodes the shot noise contamination of each mode and is related to the effective volume of the survey (Tegmark, 1997),

$$V_{\rm eff}(k) = \int {\rm d}^3x\left(\frac{\overline{n}(\boldsymbol{x})\,P(k)}{1+\overline{n}(\boldsymbol{x})\,P(k)}\right)^2, \tag{5.24}$$

such that for a constant number density, $\nu_{\rm eff}(k) = \sqrt{V_{\rm eff}(k)/V}$.

The discrete form of the LCF can now be computed by substituting Eq. (5.22) into Eq. (5.3) and proceeding as before, whereupon we see that we may write the *effective* LCF as:

$$\ell_{\rm eff}(r) = \ell^{\rm d}(r) - \ell_{\rm shot}(r)\,, \tag{5.25}$$

where the second term on the right-hand-side is a pure shot-noise term which has the form:

$$\ell_{\rm shot}(r) = 8\pi^2\left(\frac{r}{4\pi}\right)^{9/2}\int_0^{2\pi/r}{\rm d}k_1\,k_1^2\int_0^{2\pi/r}{\rm d}k_2\,k_2^2\int_{-1}^{\mu_{\rm cut}}{\rm d}\mu\sqrt{\frac{\nu_{\rm eff}(k_1)\,\nu_{\rm eff}(k_2)\,\nu_{\rm eff}(|\boldsymbol{k}_1+\boldsymbol{k}_2|)}{\hat{P}_1\,\hat{P}_2\,\hat{P}_{12}}} \times\left[\frac{1}{\overline{n}}\left(P_1+P_2+P_{12}\right)+\frac{1}{\overline{n}^2}\right]\,j_0(|\boldsymbol{k}_2-\boldsymbol{k}_1|r)\,, \tag{5.26}$$

where

$$\mu_{\rm cut} = \min\left\{1,\,\max\left\{-1,\left[(2\pi/r)^2-k_1^2-k_2^2\right]/2k_1k_2\right\}\right\} \tag{5.27}$$

guarantees that $|\boldsymbol{k}_1+\boldsymbol{k}_2| \leq 2\pi/r$ and $\hat{P}_{12} \equiv \hat{P}(|\boldsymbol{k}_1+\boldsymbol{k}_2|)$. Note that unlike $P^{\rm d}$ and $B^{\rm d}$ which may be fully corrected for 'Poisson-like' shot noise, the LCF can not (hence the name effective) since the shot noise contribution that enters via the factor $\sqrt{\nu_{\rm eff}(k_1)\ldots/P_1^{\rm d}\ldots}$ in Eq. (5.22) can not be fully separated. Hence, it remains part of the estimator, but note that in the limit of $\bar{n}P \gg 1$ the effect of shot-noise is negligible and we fully recover $\ell_{\rm m}$. On the other hand, in the limit of $\bar{n}P \ll 1$, the estimate is shot noise dominated and

the LCF scales as $\propto 1/\sqrt{\bar{n}}$ multiplied by the three-point self-correlation function for a k-space top-hat filter function evaluated for a line configuration with the regularisation factor.

The subtraction of the shot-noise terms as described above leads to a suppression of the GLCF that has to be taken into account when comparing measurements to model predictions. Figure 5.1, bottom panel, shows the impact of this suppression effect on the true dark matter LCF (thick, solid line) for two different number densities. The smallest scales are most heavily affected, which is reasonable as most of the high k-modes are efficiently damped away by $\nu_{\rm eff}$. Even for a number density of $\overline{n} = 10^{-3}\,(h^{-1}\mathrm{Mpc})^{-3}$ the suppression is significant, ranging from $\sim 70\,\%$ at $10\,h^{-1}$Mpc to still $\sim 15\,\%$ at the scale of the first BAO bump ($50\,h^{-1}$Mpc). In Sec. 5.4 we will confront Eq. (5.22) with measurements from N-body simulations.

Before moving on, we note that the suppression of the GLCF due to point sampling can be understood in a rather intuitive way: when reconstructing the matter field, each tracer contributes with a single peak convolved with some narrow window function. If the density of tracers is decreased further and further, this field will tend to look like a collection of many separate peaks of nearly equal heights instead of reflecting the true underlying matter field with its density peaks of various heights and sizes. This leads to a suppression of phase correlations because the presence of many peaks with comparable heights renders the phase distribution nearly uniform, as has already been noted above (Hikage et al., 2004).

5.3 The covariance of the line correlation

In order to study how much cosmological information a combined measurement of the LCF and power spectrum provides, we will need to compute the auto-covariance properties of the LCF and its cross-covariance with the power spectrum. The aim of this section is to provide analytic expressions for these quantities. Since this is a rather technical section, we suggest that for those not wishing to plough through the calculations at this stage they skip Secs. 5.3.2 and 5.3.3.

5.3.1 Estimators

In general we write the full joint covariance matrix of the LCF and power spectrum as:

$$\mathsf{C}_{ij} \equiv \langle \delta X_i \, \delta X_j \rangle \,, \tag{5.28}$$

where $\delta X_i = X_i - \langle X_i \rangle$ and X_i can either stand for the LCF estimator $\hat{\ell}_i$ at some radial bin r_i, or the power spectrum estimator $\hat{P}_i$ with bin k_i. In order to obtain expressions for

the estimators of the theoretical definitions in Eqs. (5.3) and (5.6) we apply the following prescription: we assume that the survey volume is large enough to encompass many independent patches of the universe and hence replace the ensemble average with an average over volume. Performing the corresponding integrations, we are able to write the estimator for the LCF as:

$$\hat{\ell}(r) = \frac{V^2}{(2\pi)^6}\left(\frac{r^3}{V}\right)^{3/2} \iint\limits_{\substack{|\boldsymbol{k}_1|,\,|\boldsymbol{k}_2|,\\ |\boldsymbol{k}_1+\boldsymbol{k}_2|\leq 2\pi/r}} \mathrm{d}^3k_1\,\mathrm{d}^3k_2\; j_0(|\boldsymbol{k}_1-\boldsymbol{k}_2|\,r)\,\epsilon_{\boldsymbol{k}_1}\,\epsilon_{\boldsymbol{k}_2}\,\epsilon_{-\boldsymbol{k}_1-\boldsymbol{k}_2}\,. \tag{5.29}$$

Similarly, for the power spectrum averaged over a bin of with Δk,

$$\hat{P}(k) = \frac{1}{V}\int_k \mathrm{d}^3q_1 \int_k \mathrm{d}^3q_2 \frac{\delta_D(\boldsymbol{q}_1+\boldsymbol{q}_2)}{V_P(k)}\,\delta_{\boldsymbol{q}_1}\,\delta_{\boldsymbol{q}_2}\,, \tag{5.30}$$

where the integrals run over $|\boldsymbol{q}| \in [k-\Delta k/2, k+\Delta k/2]$ and $V_P \equiv 4\pi k^2\Delta k$ is the volume of a spherical shell in Fourier space.

Our main task is then to evaluate ensemble averages of phase factors $\exp(i\theta_{\boldsymbol{k}})$ as well as combinations of phase factors with amplitudes $A_{\boldsymbol{k}}$, which are given as integrals over the joint PDF of Fourier modes $\mathscr{P}(\{A_{\boldsymbol{k}},\theta_{\boldsymbol{k}}\})$. App. 5.A demonstrates how this PDF can be expanded perturbatively in a series containing all higher-order spectra, where the order of the contributing terms can be conveniently labeled by powers of $1/\sqrt{V}$. This expansion is then used to derive all ensemble averages to lowest order needed for the subsequent computations.

5.3.2 Auto-covariance matrix of $\hat{\ell}$

According to Eq. (5.28), the central quantity for the LCF covariance is the six-point phase correlator subtracted by the mean,

$$\mathscr{E}_{\ell\ell} \equiv \langle \epsilon_{\boldsymbol{k}_1}\epsilon_{\boldsymbol{k}_2}\epsilon_{\boldsymbol{k}_3}\epsilon_{\boldsymbol{q}_1}\epsilon_{\boldsymbol{q}_2}\epsilon_{\boldsymbol{q}_3}\rangle - \langle \epsilon_{\boldsymbol{k}_1}\epsilon_{\boldsymbol{k}_2}\epsilon_{\boldsymbol{k}_3}\rangle\,\langle \epsilon_{\boldsymbol{q}_1}\epsilon_{\boldsymbol{q}_2}\epsilon_{\boldsymbol{q}_3}\rangle\,, \tag{5.31}$$

where it is implied that $\boldsymbol{k}_3 = -\boldsymbol{k}_1-\boldsymbol{k}_2$ and $\boldsymbol{q}_3 = -\boldsymbol{q}_1-\boldsymbol{q}_2$. Using the cumulant expansion theorem, we can split this correlator into its various connected pieces as follows,

$$\begin{aligned}\mathscr{E}_{\ell\ell} = &\Big[\langle \epsilon_{\boldsymbol{k}_1}\epsilon_{\boldsymbol{q}_1}\rangle\,\langle \epsilon_{\boldsymbol{k}_2}\epsilon_{\boldsymbol{q}_2}\rangle\,\langle \epsilon_{\boldsymbol{k}_3}\epsilon_{\boldsymbol{q}_3}\rangle + \text{sym.(6)}\Big] + \Big[\langle \epsilon_{\boldsymbol{q}_1}\epsilon_{\boldsymbol{k}_2}\epsilon_{\boldsymbol{k}_3}\rangle_c\,\langle \epsilon_{\boldsymbol{k}_1}\epsilon_{\boldsymbol{q}_2}\epsilon_{\boldsymbol{q}_3}\rangle_c + \text{sym.(9)}\Big] \\ &+ \Big[\langle \epsilon_{\boldsymbol{k}_1}\epsilon_{\boldsymbol{k}_2}\epsilon_{\boldsymbol{q}_1}\epsilon_{\boldsymbol{q}_2}\rangle_c\,\langle \epsilon_{\boldsymbol{k}_3}\epsilon_{\boldsymbol{q}_3}\rangle + \text{sym.(9)}\Big] + \langle \epsilon_{\boldsymbol{k}_1}\epsilon_{\boldsymbol{k}_2}\epsilon_{\boldsymbol{k}_3}\epsilon_{\boldsymbol{q}_1}\epsilon_{\boldsymbol{q}_2}\epsilon_{\boldsymbol{q}_3}\rangle_c\,,\end{aligned} \tag{5.32}$$

with sym.(n) indicating that $n-1$ terms have to be added to symmetrize the corresponding expressions with respect to the $\boldsymbol{k}$'s and $\boldsymbol{q}$'s. Connected correlators consist of all those terms that cannot be written as a product of two or more connected pieces and for the phase fields they result from the PDF expansion coefficients where $m=1$ in Eq. (5.84).

As such, each connected phase correlator is not limited to a single spectrum but a series of all even or odd spectra, which become increasingly suppressed by factors of volume. The exception is the two-point function of phase factors, where statistical homogeneity dictates that $\boldsymbol{k}_1 = -\boldsymbol{k}_2$, in which case $\epsilon_{\boldsymbol{k}_1}\epsilon_{\boldsymbol{k}_2} = |\epsilon_{\boldsymbol{k}_1}|^2$, which is strictly one. Hence,

$$\langle \epsilon_{\boldsymbol{k}_1}\epsilon_{\boldsymbol{k}_2}\rangle = \frac{(2\pi)^3}{V}\delta_D(\boldsymbol{k}_1+\boldsymbol{k}_2)\,, \tag{5.33}$$

and we see that pairing either two $\boldsymbol{k}$- or $\boldsymbol{q}$-modes in $\mathscr{E}_{\ell\ell}$ causes the respective third wavevector to be zero, meaning that it must belong to the background. These modes do not contribute to the correlation functions and consequently all those terms were left out in Eq. (5.32).

Since one of the three Dirac delta functions is redundant and $\delta_D(\boldsymbol{0}) = V/(2\pi)^3$, the Gaussian part of Eq. (5.32) is of fourth order, i.e. of order $1/V^2$. The second line is a product of two three-point correlators, which are given by Eq. (5.5) and by eliminating one of the two delta functions we have,

$$\langle \epsilon_{\boldsymbol{q}_1}\epsilon_{\boldsymbol{k}_2}\epsilon_{\boldsymbol{k}_3}\rangle_c \langle \epsilon_{\boldsymbol{k}_1}\epsilon_{\boldsymbol{q}_2}\epsilon_{\boldsymbol{q}_3}\rangle_c = \frac{(2\pi)^3}{V}\left(\frac{\sqrt{\pi}}{2}\right)^6 \delta_D(\boldsymbol{k}_1+\boldsymbol{q}_1)\, p^{(3)}(\boldsymbol{q}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)\, p^{(3)}(\boldsymbol{k}_1,\boldsymbol{q}_2,\boldsymbol{q}_3)\,. \tag{5.34}$$

The quantities $p^{(N)}$ refer to the reduced N-th order spectra, which are defined in Eq. (5.82) and have a volume dependence of $\propto V^{N/2-1}$. This implies that the expression above is of the same order as the Gaussian term, such that higher-order corrections to the three-point correlator do not have to be taken into account. The next two contributions to $\mathscr{E}_{\ell\ell}$ contain connected four- and six-point correlators, which are worked out in App. 5.A.1 and 5.A.2, giving the following results to lowest order,

$$\langle \epsilon_{\boldsymbol{q}_1}\epsilon_{\boldsymbol{k}_2}\epsilon_{\boldsymbol{q}_1}\epsilon_{\boldsymbol{q}_2}\rangle_c \langle \epsilon_{\boldsymbol{k}_3}\epsilon_{\boldsymbol{q}_3}\rangle = \frac{(2\pi)^3}{V}\left(\frac{\sqrt{\pi}}{2}\right)^4 \delta_D(\boldsymbol{k}_3+\boldsymbol{q}_3)\, p^{(4)}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{q}_1,\boldsymbol{q}_2)\,, \tag{5.35}$$

$$\langle \epsilon_{\boldsymbol{k}_1}\epsilon_{\boldsymbol{k}_2}\epsilon_{\boldsymbol{k}_3}\epsilon_{\boldsymbol{q}_1}\epsilon_{\boldsymbol{q}_2}\epsilon_{\boldsymbol{q}_3}\rangle_c = \left(\frac{\sqrt{\pi}}{2}\right)^6 p^{(6)}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{q}_1,\boldsymbol{q}_2,\boldsymbol{q}_3)\,. \tag{5.36}$$

Having all necessary ingredients, we can plug the phase correlators back into Eq. (5.28) and perform the trivial Fourier integrations over the delta functions. Summing up all contributions in Eq. (5.32), it is possible to show that

$$\langle \delta\hat{\ell}_i\, \delta\hat{\ell}_j\rangle = \left[\frac{(r_i\, r_j)^3}{V^2}\right]^{3/2} \left\{\mathscr{C}_G + \mathscr{C}_T + \mathscr{C}_{B^2} + \mathscr{C}_{P_6}\right\}, \tag{5.37}$$

with

$$\mathscr{C}_G \equiv \iint\limits_{\substack{|\boldsymbol{k}_1|,\,|\boldsymbol{k}_2|,\\ |\boldsymbol{k}_1+\boldsymbol{k}_2|\,\leq\,2\pi/r}} \frac{\mathrm{d}^3k_1}{k_f^3}\,\frac{\mathrm{d}^3k_2}{k_f^3}\left[j_0(|2\boldsymbol{k}_1+\boldsymbol{k}_2|\,r_i)\,\mathscr{J}(\boldsymbol{k}_1,\boldsymbol{k}_2,r_j)+(r_i\leftrightarrow r_j)\right], \tag{5.38}$$

$$\mathscr{C}_T \equiv \left(\frac{\sqrt{\pi}}{2}\right)^4 \int\limits_{|\boldsymbol{q}|\,\leq\,2\pi/r} \frac{\mathrm{d}^3q}{k_f^3} \int\limits_{\substack{|\boldsymbol{k}_1|,\,|\boldsymbol{k}_1+\boldsymbol{q}|\\ \leq\,2\pi/r_i}} \frac{\mathrm{d}^3k_1}{k_f^3} \int\limits_{\substack{|\boldsymbol{k}_2|,\,|\boldsymbol{k}_2+\boldsymbol{q}|\\ \leq\,2\pi/r_j}} \frac{\mathrm{d}^3k_2}{k_f^3}\,\mathscr{J}(\boldsymbol{q},\boldsymbol{k}_1,r_i)\,\mathscr{J}(\boldsymbol{q},\boldsymbol{k}_2,r_j)$$
$$\times\; p^{(4)}(\boldsymbol{q},\boldsymbol{k}_1,\boldsymbol{k}_2,-\boldsymbol{q}-\boldsymbol{k}_1-\boldsymbol{k}_2)\,, \tag{5.39}$$

$$\mathscr{C}_{B^2} \equiv \left(\frac{\sqrt{\pi}}{2}\right)^6 \int\limits_{|\boldsymbol{q}|\,\leq\,2\pi/r} \frac{\mathrm{d}^3q}{k_f^3}\left[\left(\int\limits_{\substack{|\boldsymbol{k}_1|,\,|\boldsymbol{k}_1+\boldsymbol{q}|\\ \leq\,2\pi/r_i}} \frac{\mathrm{d}^3k_1}{k_f^3}\,\mathscr{J}(\boldsymbol{q},\boldsymbol{k}_1,r_i)\,p^{(3)}(\boldsymbol{q},\boldsymbol{k}_1,-\boldsymbol{q}-\boldsymbol{k}_1)\right)\times\left(r_i\leftrightarrow r_j\right)\right], \tag{5.40}$$

$$\mathscr{C}_{P_6} \equiv \left(\frac{\sqrt{\pi}}{2}\right)^6 \iint\limits_{\substack{|\boldsymbol{k}_1|,\,|\boldsymbol{k}_2|,\\ |\boldsymbol{k}_1+\boldsymbol{k}_2|\,\leq\,2\pi/r_i}} \frac{\mathrm{d}^3k_1}{k_f^3}\,\frac{\mathrm{d}^3k_2}{k_f^3} \iint\limits_{\substack{|\boldsymbol{q}_1|,\,|\boldsymbol{q}_2|,\\ |\boldsymbol{q}_1+\boldsymbol{q}_2|\,\leq\,2\pi/r_j}} \frac{\mathrm{d}^3q_1}{k_f^3}\,\frac{\mathrm{d}^3q_2}{k_f^3}$$
$$\times\; j_0(|\boldsymbol{k}_1-\boldsymbol{k}_2|\,r_i)\,j_0(|\boldsymbol{q}_1-\boldsymbol{q}_2|\,r_j)\,p^{(6)}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{q}_1,\boldsymbol{q}_2,\boldsymbol{q}_3)\,, \tag{5.41}$$

where we have defined the kernel function

$$\mathscr{J}(\boldsymbol{k}_1,\boldsymbol{k}_2,r) \equiv 2\,j_0(|\boldsymbol{k}_1-\boldsymbol{k}_2|\,r)+j_0(|2\boldsymbol{k}_1+\boldsymbol{k}_2|\,r)\,. \tag{5.42}$$

Furthermore, $k_f \equiv 2\pi/V^{\frac{1}{3}}$ denotes the fundamental frequency and $r \equiv \max\{r_i,\,r_j\}$. We note that Eq. (5.37) closely resembles the covariance of the bispectrum (see Sefusatti et al., 2006) with a Gaussian term, one that is quadratic in the bispectrum as well as terms proportional to the trispectrum and the sixth-order spectrum. However, one important difference is that there is no cosmology dependence in the Gaussian contribution, which is a direct consequence of the two-point phase correlator in Eq. (5.33), that carries no information either. That means that Eq. (5.38) is just an algebraic expression depending on the scales r_i and r_j, and for the variance we get explicitely,

$$\mathrm{Var}\left(\hat{\ell}(r)\right) \approx 0.25\,\frac{r^3}{V}\,. \tag{5.43}$$

5.3.3 Cross-covariance matrix between $\hat{P}$ and $\hat{\ell}$

Analogously to the last section we can now derive the cross-covariance between LCF and power spectrum. In this case we deal with a mixed five-point correlator of phase factors and amplitudes,

$$\mathscr{E}_{P\ell} \equiv \langle \delta_{\boldsymbol{k}_1} \delta_{\boldsymbol{k}_2} \epsilon_{\boldsymbol{q}_1} \epsilon_{\boldsymbol{q}_2} \epsilon_{\boldsymbol{q}_3} \rangle - \langle \delta_{\boldsymbol{k}_1} \delta_{\boldsymbol{k}_2} \rangle \langle \epsilon_{\boldsymbol{q}_1} \epsilon_{\boldsymbol{q}_2} \epsilon_{\boldsymbol{q}_3} \rangle \,. \tag{5.44}$$

Let us again begin by splitting this expression into its connected correlators, giving

$$\mathscr{E}_{P\ell} = \Big[\langle \delta_{\boldsymbol{k}_1} \epsilon_{\boldsymbol{q}_1} \rangle \langle \delta_{\boldsymbol{k}_2} \epsilon_{\boldsymbol{q}_2} \epsilon_{\boldsymbol{q}_3} \rangle_c + \text{sym.(3)} \Big] + \big(\boldsymbol{k}_1 \leftrightarrow \boldsymbol{k}_2 \big) + \langle \delta_{\boldsymbol{k}_1} \delta_{\boldsymbol{k}_2} \epsilon_{\boldsymbol{q}_1} \epsilon_{\boldsymbol{q}_2} \epsilon_{\boldsymbol{q}_3} \rangle_c \,, \tag{5.45}$$

where we have left out all terms that give rise to background modes as before. Due to statistical homogeneity the mixed two-point correlator is simply the average of the amplitude $|\delta_{\boldsymbol{k}}|$, which must be evaluated using the joint PDF of Fourier modes described in App. 5.A. Assuming temporarily that we are having a discrete set of modes, the Gaussian part is given by (see Eqs. 5.80 and 5.87)

$$\begin{aligned} \langle \delta_{\boldsymbol{k}} \epsilon_{\boldsymbol{q}} \rangle &= \sqrt{V P(k)} \int \prod_{\boldsymbol{p} \in \text{uhs}} 2 A_{\boldsymbol{p}} \mathrm{e}^{-A_{\boldsymbol{p}}^2} \, \mathrm{d}A_{\boldsymbol{p}} \, \frac{\mathrm{d}\theta_{\boldsymbol{p}}}{2\pi} \, A_{\boldsymbol{k}} \, \mathrm{e}^{i(\theta_{\boldsymbol{k}} + \theta_{\boldsymbol{q}})} \\ &= \frac{\sqrt{\pi}}{2} \sqrt{V P(k)} \int \prod_{\boldsymbol{p} \in \text{uhs}} \frac{\mathrm{d}\theta_{\boldsymbol{p}}}{2\pi} \, \mathrm{e}^{i(\theta_{\boldsymbol{k}} + \theta_{\boldsymbol{q}})} \\ &= \frac{\sqrt{\pi}}{2} \sqrt{V P(k)} \, \delta^K_{\boldsymbol{k}+\boldsymbol{q}} \,, \end{aligned} \tag{5.46}$$

where the products run over all modes $\boldsymbol{p}$ in the upper half sphere (uhs), defined by $p_z \geq 0$. In going from the first to the second line we have made use of Eq. (5.91) to do the integrals over $A_{\boldsymbol{p}}$ and then we see that the remaining integrals only give a non-vanishing result if the two phases cancel out each other. The square root factor of the power spectrum enters because of our choice of normalization, i.e. $|\delta_{\boldsymbol{k}}| \equiv \sqrt{V P(k)} \, A_{\boldsymbol{k}}$. Finally, taking the continuum limit,

$$\langle \delta_{\boldsymbol{k}} \epsilon_{\boldsymbol{q}} \rangle = \frac{(2\pi)^3}{V} \frac{\sqrt{\pi}}{2} \sqrt{V P(k)} \, \delta_D(\boldsymbol{k} + \boldsymbol{q}) \,. \tag{5.47}$$

In a similar manner we can compute the mixed three-point correlator and we consider the case where $\boldsymbol{q}_1, \boldsymbol{q}_2 \in$ uhs, while $\boldsymbol{k} \in$ lhs. The lowest-order term that is contributing is proportional to the reduced bispectrum and we have,

$$\begin{aligned} \langle \delta_{\boldsymbol{k}} \epsilon_{\boldsymbol{q}_1} \epsilon_{\boldsymbol{q}_2} \rangle_c = \sqrt{V P(k)} \int \prod_{\boldsymbol{p} \in \text{uhs}} 2 A_{\boldsymbol{p}} \mathrm{e}^{-A_{\boldsymbol{p}}^2} \, \mathrm{d}A_{\boldsymbol{p}} \, \frac{\mathrm{d}\theta_{\boldsymbol{p}}}{2\pi} \sum_{\substack{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3 \in \text{uhs} \\ \boldsymbol{u}_i \neq \boldsymbol{u}_j}} A_{\boldsymbol{u}_1} A_{\boldsymbol{u}_2} A_{\boldsymbol{u}_3} \\ \times \, p^{(3)}(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3) \, A_{\boldsymbol{k}} \cos\big(\theta_{\boldsymbol{u}_1} + \theta_{\boldsymbol{u}_2} - \theta_{\boldsymbol{u}_3}\big) \, \mathrm{e}^{i(\theta_{\boldsymbol{q}_1} + \theta_{\boldsymbol{q}_2} - \theta_{\boldsymbol{k}})} \,. \end{aligned} \tag{5.48}$$

The phase factors can only be fully cancelled in the case where $\boldsymbol{k}$ equals $\boldsymbol{u}_3$, such that the integrals over $A_{\boldsymbol{p}}$ produce a factor of $(\sqrt{\pi}/2)^2$. For the remaining phase integrals we need to impose the condition $\boldsymbol{k}+\boldsymbol{q}_1+\boldsymbol{q}_2=0$ and hence, the final result after taking the continuum limit is

$$\langle \delta_{\boldsymbol{k}}\epsilon_{\boldsymbol{q}_1}\epsilon_{\boldsymbol{q}_2}\rangle_c = \frac{(2\pi)^3}{V}\left(\frac{\sqrt{\pi}}{2}\right)^2 \sqrt{V\,P(k)}\,p^{(3)}(\boldsymbol{k},\boldsymbol{q}_1,\boldsymbol{q}_2)\,\delta_D(\boldsymbol{k}+\boldsymbol{q}_1+\boldsymbol{q}_2)\,. \tag{5.49}$$

One can check that this result holds for any configuration of the three vectors $\boldsymbol{k}$, $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$. The calculation of the connected five-point correlator is more cumbersome and is therefore carried out in App. 5.A.3, giving the simple outcome

$$\begin{aligned}\langle \delta_{\boldsymbol{k}_1}\delta_{\boldsymbol{k}_2}\epsilon_{\boldsymbol{q}_1}\epsilon_{\boldsymbol{q}_2}\epsilon_{\boldsymbol{q}_3}\rangle_c = (2\pi)^3\left(\frac{\sqrt{\pi}}{2}\right)^3 \sqrt{P(k_1)\,P(k_2)}\,p^{(5)}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{q}_1,\boldsymbol{q}_2,\boldsymbol{q}_3)\\ \times\,\delta_D(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{q}_1+\boldsymbol{q}_2+\boldsymbol{q}_3)\,.\end{aligned} \tag{5.50}$$

Finally, assembling all these pieces in Eq. (5.45) and plugging back into Eq. (5.28) we obtain,

$$\langle \delta\hat{\ell}_i\,\delta\hat{P}_j\rangle = \frac{1}{k_f^3}\left(\frac{\sqrt{\pi}}{2}\right)^3\left(\frac{r_i^3}{V}\right)^{3/2}\left[\mathscr{C}_{PB}+\mathscr{C}_{P_5}\right], \tag{5.51}$$

with

$$\mathscr{C}_{PB} \equiv 2\Theta_{ij}\,P(k_j)\int\limits_{|\boldsymbol{q}|,\,|\boldsymbol{q}+\boldsymbol{k}_j|\,\leq\,2\pi/r_i} \mathrm{d}^3q\; j_0(|\boldsymbol{k}_j-\boldsymbol{q}|r_i)\,p^{(3)}(\boldsymbol{k}_j,\boldsymbol{q},-\boldsymbol{k}_j-\boldsymbol{q})\,, \tag{5.52}$$

and

$$\mathscr{C}_{P_5} \equiv \frac{1}{V_P(k_j)}\int_{k_j}\mathrm{d}^3p \iint\limits_{\substack{|\boldsymbol{q}_1|,\,|\boldsymbol{q}_2|,\\ |\boldsymbol{q}_1+\boldsymbol{q}_2|\,\leq\,2\pi/r_i}} \mathrm{d}^3q_1\,\mathrm{d}^3q_2\; j_0(|\boldsymbol{q}_1-\boldsymbol{q}_2|r_i)\,P(p)\,p^{(5)}(\boldsymbol{p},-\boldsymbol{p},\boldsymbol{q}_1,\boldsymbol{q}_2,\boldsymbol{q}_3)\,. \tag{5.53}$$

Here, Θ_{ij} stands for the theta function $\Theta(1-k_j\,r_i/2\pi)$, indicating that we only get a correlation at the order of the bispectrum if the power spectrum scale k_j lies outside of the region affected by the LCF cutoff. As for the LCF auto-covariance we notice a strong similarity to the bispectrum and power spectrum cross-covariance given in Sefusatti et al. (2006).

5.3.4 Signal-to-noise

In summary, we see that the LCF covariance closely resembles the covariance of the bispectrum. Its leading-order term is much simpler, though, as it carries no cosmological information – this enters only through higher-order terms. In the large-scale limit the LCF covariance is therefore independent of redshift and shot noise. On the other hand the lowest-order contribution to the cross-covariance between $\hat{P}$ and $\hat{\ell}$ contains cosmological information from both, the power spectrum and the bispectrum.

Based on the results from the previous section, it is instructive to compare the cumulative signal-to-noise of the LCF with that of the power spectrum in the Gaussian approximation. To plot both as a function of the maximal mode $k_{\rm max}$ included, we use the correspondence $r = \pi/k$, so that we can write the LCF signal-to-noise as follows

$$\left(\frac{\mathscr{S}}{\mathscr{N}}\right)^2_{\ell} = \sum_{i=1}^{i_{\rm max}} \sum_{j=1}^{i_{\rm max}} \ell_{\rm eff}(r_i)\, \mathsf{C}^{-1}_{ij}\, \ell_{\rm eff}(r_j)\,, \tag{5.54}$$

where $r_i = \pi/(i\,\delta k)$ and $i_{\rm max} = k_{\rm max}/\delta k$. The signal $\ell_{\rm eff}(r)$ denotes the discrete LCF subtracted by the shot noise terms that appear in the square bracket of Eq. (5.22) and hence takes the suppression due to $\nu_{\rm eff}$ into account (explicit expressions are given in Eqs. 5.25 and 5.26).

The Gaussian part of the covariance matrix for the power spectrum estimator in Eq. (5.30) is diagonal and given by Feldman et al. (1994)

$$\sigma^2_P(k) = \frac{2(2\pi)^3}{V_P(k)\,V}\left(P(k) + \frac{1}{\overline{n}}\right)^2 = \frac{2(2\pi)^3}{V_P(k)}\,\frac{P(k)}{V_{\rm eff}(k)}\,, \tag{5.55}$$

with $P(k)$ meaning the discrete power spectrum subtracted by its shot noise component and $V_{\rm eff}$ the effective volume introduced in Eq. (5.24). Writing $k = i\,\delta k$ the signal-to-noise is thus

$$\left(\frac{\mathscr{S}}{\mathscr{N}}\right)^2_{P} = \sum_{i=1}^{i_{\rm max}} \frac{P(k)^2}{\sigma^2_P(k)} = \sum_{i=1}^{i_{\rm max}} \frac{V_P(k)}{2(2\pi)^3}\, V_{\rm eff}(k)\,. \tag{5.56}$$

In Fig. 5.2 we plot Eqs. (5.54) and (5.56) for various number densities as a function of $k_{\rm max}\,(r_{\rm min})$, where we assumed a cubical survey volume with sidelength $L = 1.5\,h^{-1}\,$Gpc and used a bin width of $\delta k = 2\pi/L$. As before, we adopted a ΛCDM cosmology with parameters described in Sec. 5.4 and computed the LCF in linear PT. We see that, for most of the scales that we have considered, $k_{\rm max} < 0.3\,h\,{\rm Mpc}^{-1}$, the power spectrum signal-to-noise dominates over that of the LCF, being comparable only for $k_{\rm max} \sim 0.3\,h\,{\rm Mpc}^{-1}$ and the densest galaxy sample. However, the LCF signal-to-noise increases more quickly as a function of $k_{\rm max}$, which is a recognized feature of higher-order statistics (Sefusatti and Scoccimarro, 2005). On the other hand, the LCF is also more heavily impacted

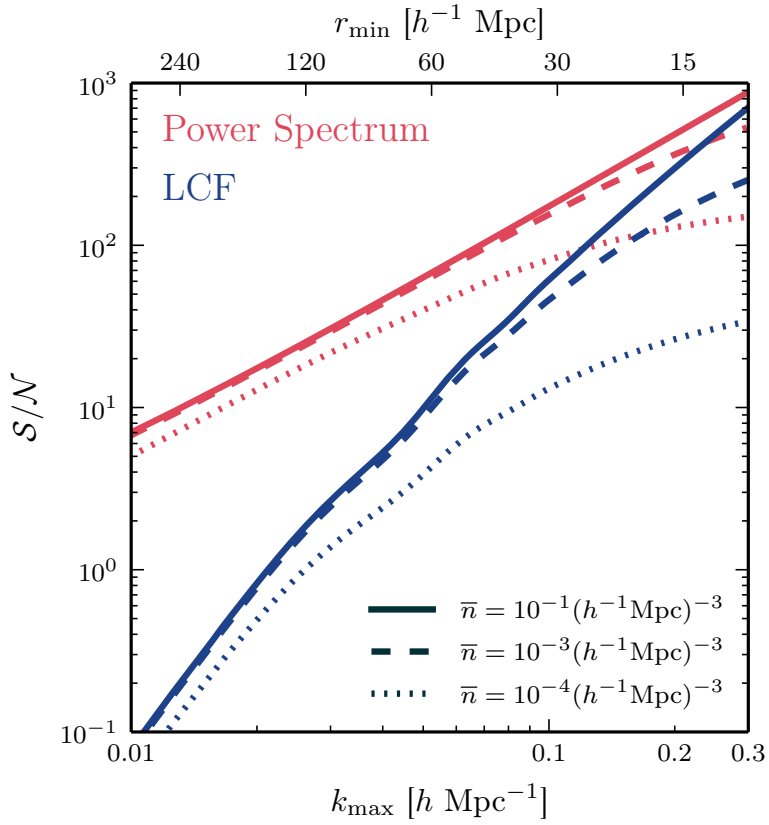


Fig. 5.2 Cumulative signal-to-noise for the power spectrum (red) and LCF (blue) based on Eqs. (5.56) and (5.54) as a function of the maximal mode $k_{\rm max}$ or minimal scale $r_{\rm min}$, which are related via $r_{\rm min} = \pi/k_{\rm max}$. We consider three different galaxy number densities: $\overline{n} = 10^{-3}\,(h^{-1}\,\mathrm{Mpc})^{-3}$ (solid lines), $10^{-4}\,(h^{-1}\,\mathrm{Mpc})^{-3}$ (dashed) and $10^{-5}\,(h^{-1}\,\mathrm{Mpc})^{-3}$ (dotted).

by shot noise than the power spectrum. While the difference in signal-to-noise at $k_{\rm max} = 0.3\,h\,\mathrm{Mpc}^{-1}$ is $\sim 10\,\%$ for the highest number density it is already $\sim 75\,\%$ for the lowest. This is because each mode contributing to the signal-to-noise is penalized by a factor of $\nu_{\rm eff}$, so for an N-th order statistic we should expect a suppression proportional to $\nu_{\rm eff}^N$.

Finally, we note that here we are only probing a subset of the available information in the three-point phase correlation as the definition of the LCF in Eq. (5.3) forces the three points of the triangle to lie along a line – so called degenerate triangles. Adding measurements with various other triangle configurations is certainly going to increase the signal-to-noise. Another possibility to boost the signal-to-noise is to find a more optimal mode cutoff than the top-hat window used in the definition of the LCF, which could for instance be done in the Gaussian approximation of Eq. (5.37).

5.4 Comparison with N-body simulations

5.4.1 Numerical simulations

In this section we present measurements of the LCF and its covariance matrix along with the cross-correlation of the LCF with the power spectrum. These measurements are based on a set of 200 dark matter only N-body simulations, which were run on the ZBOX2 and ZBOX3 supercomputers at the University of Zurich (see Sec. 6 of Smith, 2009) using the `Gadget-2` code of Springel (2005). The simulations contain 750^3 particles, which are enclosed in a periodic box of comoving size $L = 1500\,h^{-1}\,\mathrm{Mpc}$. Initial conditions were set up at redshift $z = 49$ based on different realizations of a Gaussian random field and second-order Lagrangian perturbation theory (Crocce and Scoccimarro, 2006b) for the displacement of the particles. The power spectrum of the Gaussian random fields was determined from a transfer function generated by `CMBFAST` (Seljak and Zaldarriaga, 1996) assuming a flat ΛCDM model with cosmological parameters $\Omega_m = 0.25$, $\Omega_b = 0.04$, $\sigma_8 = 0.8$, $n_s = 1.0$ and $h = 0.7$.

5.4.2 Estimating $\hat{P}$ in simulations

From the simulations we construct smooth dark matter density fields by distributing the particles onto a grid using a cloud-in-cell (CIC) assignment scheme with $N = 512$ cells per side. Each Fourier mode of the resulting field is then corrected for the convolution with the mesh by dividing out the Fourier transform of the CIC window function:

$$\delta^d_{\boldsymbol{k}} = \frac{\delta^g_{\boldsymbol{k}}}{W_{\mathrm{CIC}}(\boldsymbol{k})}\,, \tag{5.57}$$

where

$$W_{\mathrm{CIC}}(\boldsymbol{k}) = \prod_{i=1}^{3}\left[\frac{\sin\left(\pi k_i/2k_{\mathrm{Ny}}\right)}{\pi k_i/2k_{\mathrm{Ny}}}\right]^2\,. \tag{5.58}$$

The superscripts d and g denote discrete and grid quantities, respectively, and $k_{\mathrm{Ny}} = \pi N_{\mathrm{grid}}/L$ is the Nyquist frequency of the mesh, with N_{grid} being the number of mesh-cells per dimension.

For the power spectrum estimator presented in Eq. (5.30) it can be shown that for a finite periodic volume it can be rewritten for a given scale k as:

$$\hat{P}^d(k) = \frac{V}{N_k}\sum_{|k-q_i|\,\leq\,\Delta k/2}\left|\delta^d_{\boldsymbol{q}_i}\right|^2\,, \tag{5.59}$$

where the sum extends over all modes within a shell of thickness Δk centered around k, and where N_k is the number of Fourier modes in each shell. This estimate still

suffers from discreteness effects, and at late times when the initial transients are small, an unbiased estimate is therefore obtained by subtracting the shot noise term, i.e. $\hat{P} = \hat{P}^d - 1/\bar{n}$, where $\bar{n} = N/V$, with N being the number of dark matter particles. In the following we consider measurements of the power spectrum in 30 bins from $k_{\rm min} = 0.005$ till $k_{\rm max} = 0.3\,h\,{\rm Mpc}^{-1}$ and a bin width of $\Delta k = 0.01\,h\,{\rm Mpc}^{-1}$. The power spectrum is also susceptible to aliasing effects, hence we also use FFT grids for which $k_{\rm Ny} > 2k_{\rm max}$.

5.4.3 Estimating $\hat{\ell}$ in simulations

The discretized version of the LCF estimator given by Eq. (5.29) can be written in the form:

$$\text{E1}: \;\; \hat{\ell}^d(r) = \left(\frac{r^3}{V}\right)^{3/2} \sum_{\substack{|\boldsymbol{k}_1|,\,|\boldsymbol{k}_2|,\\ |\boldsymbol{k}_1+\boldsymbol{k}_2| \le 2\pi/r}} \overline{j_0}(|\boldsymbol{k}_1 - \boldsymbol{k}_2|r)\, \epsilon^d_{\boldsymbol{k}_1}\, \epsilon^d_{\boldsymbol{k}_2}\, \epsilon^d_{-\boldsymbol{k}_1-\boldsymbol{k}_2}\,, \tag{5.60}$$

where $\overline{j_0}(|\boldsymbol{k}|r)$ denotes the spherical Bessel function averaged over the k-space volume centred on the Fourier mode $\boldsymbol{k}$. We found that this estimator is computationally expensive to estimate, at least for the case where r is probing small scales, since the 6D sum in Eq. (5.60) runs over the majority of Fourier modes – the worst case being $\mathcal{O}\left(N^6_{\rm grid}\right)$ terms.

In order to accelerate the estimation we employ an implementation based on the real space phase fields. The estimator is built around Eq. (5.1): we take the product of the $\epsilon_r(\boldsymbol{x})$ values at three different points separated by scale r and average these over all possible positions and orientations of the three points in a line (see also App. 4.A). The new estimator can be written:

$$\text{E2}: \;\; \hat{\ell}^d(r) = \left(\frac{r^3}{V}\right)^{3/2} \frac{\Delta\varphi\,\Delta\vartheta}{16\pi} \sum_{\boldsymbol{x}} \sum_{i,j=0}^{n} w_i\, w_j\, \sin\left(j\,\Delta\vartheta\right) \epsilon^d_r(\boldsymbol{x})\, \epsilon^d_r\left(\boldsymbol{x} + \boldsymbol{r}_{ij}\right)\, \epsilon^d_r\left(\boldsymbol{x} - \boldsymbol{r}_{ij}\right), \tag{5.61}$$

where the sum over $\boldsymbol{x}$ averages over all points in the volume, and the sums over i and j discretize the angular integration of the orientation of the line $\hat{\boldsymbol{r}}$ over all orientations, with i labeling the azimuthal angle and j the angle with respect to the polar axis. The angular bin sizes are $\Delta\varphi = 2\pi/n$ and $\Delta\vartheta = \pi/n$. The weight factors come from the trapezoidal rule for numerical integration and are either $w_i = 1$ if $i = 0$ or n and 2 otherwise. The radial vector is defined by $\boldsymbol{r}_{ij} \equiv r\,\boldsymbol{e}_r(i,j)$, where the unit vector is specified by

$$\boldsymbol{e}_r(i,j) \equiv \begin{pmatrix} \sin\vartheta\cos\varphi \\ \sin\vartheta\sin\varphi \\ \cos\vartheta \end{pmatrix} = \begin{pmatrix} \sin(j\,\Delta\vartheta)\cos(i\,\Delta\varphi) \\ \sin(j\,\Delta\vartheta)\sin(i\,\Delta\varphi) \\ \cos(j\,\Delta\vartheta) \end{pmatrix}. \tag{5.62}$$

Note that owing to the fact that the phase field is smoothed on scale r (see Eq. 5.2), $\epsilon_r^d(\boldsymbol{x})$ has to be recomputed for each new scale for which ℓ is estimated. This however can be rapidly performed by applying the cutoff in Fourier space and executing an inverse Fast Fourier Transform, i.e.

$$\epsilon_r^d(\boldsymbol{x}) = \mathtt{iFFT}\left[\epsilon_{\boldsymbol{k}}^d \Theta\left(1 - \frac{k\,r}{2\pi}\right)\right]. \tag{5.63}$$

Lastly, for the case of estimation in the N-body simulations, any point that gets mapped outside of the box is placed back according to the periodic boundary conditions.

From a number of tests we have found that the numerical estimator above converges already for a moderate number of bins, which is of the order 10. That means E2 requires of the order $100\,N^3$ operations, independent of the scale r, while E1 scales as $\sim (2L/r)^6$ after the mode cutoff has been taken into account. This implies that E2 quickly becomes more efficient at scales $r \lesssim 10^{-1/3} \times 2L/\sqrt{N}$, i.e. $r \lesssim 60\,h^{-1}$Mpc for $N = 512$.

5.4.4 Comparison of estimators with simulations

Figure 5.3, upper panel, shows the results for the matter LCF measured for three different redshifts $z = 0$, 0.52 and 1. The measurements were made for 30 bins with the line scale varying $r \in [10, 200]\ h^{-1}$Mpc, which on using the relation $k \sim \pi/r$ roughly corresponds to a wavemode range of $k \in [0.016, 0.31]\ h\,\mathrm{Mpc}^{-1}$ and is hence comparable to our power spectrum measurements. The spacing of the first 7 bins is $2.5\,h^{-1}$Mpc, increasing to $5\,h^{-1}$Mpc for the next 11 and the remaining 12 bins have a spacing of $10\,h^{-1}$Mpc. The error bars show the expected variations between realisations and are obtained from the 200 realisations. For the sake of clarity we only show the 1σ error bars for the sample with the lowest redshift. The estimates were corrected for shot noise as discussed in Sec. 5.2.4. From the figure we see immediately how the LCF increases with time, which is a result of growing phase correlations under the influence of non-linear gravity. For the same reason the LCF is a function that is mostly decreasing with scale, as at larger r the density field is in a more linear state with phases being increasingly random. The solid lines in Fig. 5.3 show the predictions from tree-level SPT (see Eq. 5.12), and on large scales ($r > 30\,h^{-1}$Mpc) are in good agreement with the data for the three redshifts considered. However, on smaller scales we see that there are departures from this lowest-order prediction, which consistently underpredicts the measured LCF.

Figure 5.3, lower panel, shows the difference $\Delta\ell = \hat{\ell} - \ell$ between the measurement and model predictions, normalized by the standard deviation. For all scales above $\sim 30\,h^{-1}$Mpc the difference is within the 1σ interval, while for smaller scales the agreement breaks down quickly, being already worse than 4σ at $r = 20\,h^{-1}$Mpc and the lowest redshift. These deviations arise because non-linear corrections to the bispectrum and power spectrum become increasingly important on these small scales. This

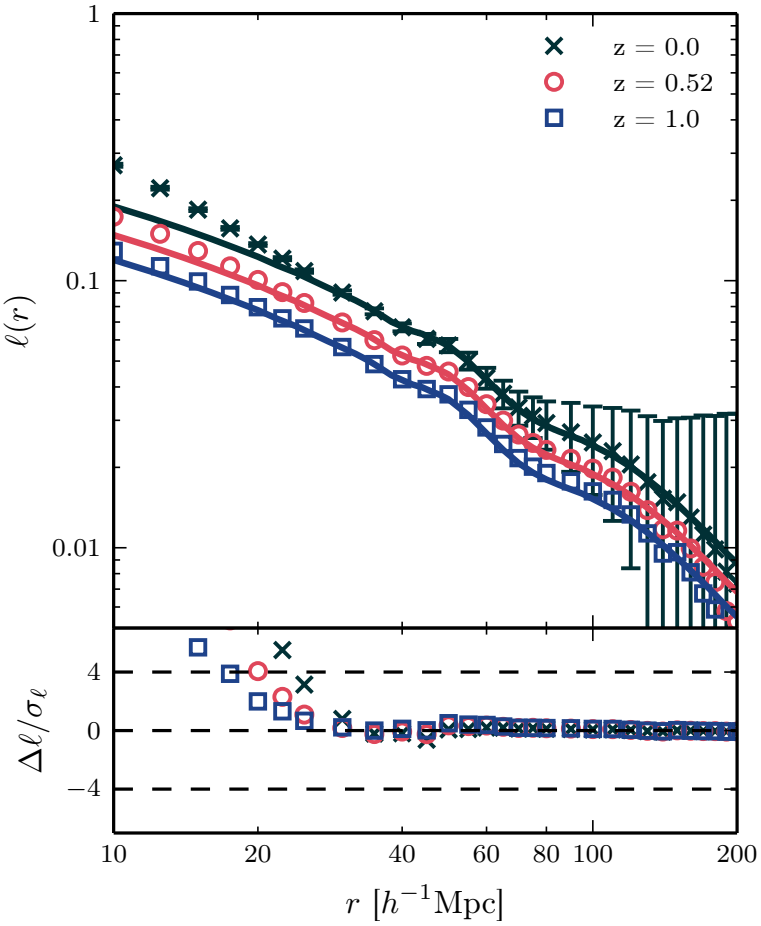


Fig. 5.3 The top panel displays the results of the LCF measurements from 200 N-body simulations (data points) for the three redshifts $z = 0$, 0.52 and 1, while the 1σ error bars for just shown for the lowest redshift sample. Solid lines of matching colour are the corresponding predictions from tree-level perturbation theory. The bottom panel shows the relative difference between measured and predicted LCF, normalised by the 1σ standard deviation.

also explains why the discrepancies are less significant for higher redshifts. Changing the power spectrum model that enters Eq. (5.12) from the linear spectrum to the power spectrum with corrections up to the one-loop level does not bring any improvement. On the contrary, we note that the increase in small-scale power leads to a further suppression of the LCF and alters the predictions by $1.4\,\sigma_\ell$ at $r = 40\,h^{-1}$Mpc. This suppression will be countered when using the appropriate one-loop bispectrum, but this seems to indicate that linear theory is applicable throughout a larger range of scales for the LCF than it is for the conventional statistical measures.

5.4.5 Testing the effects of shot-noise on $\hat{\ell}$

In the measurements above where all the dark matter particles have been used to construct the density field, shot noise has no discernible effect. To test the suppression from discreteness derived in Sec. 5.2.4 we therefore carry out the same measurements,

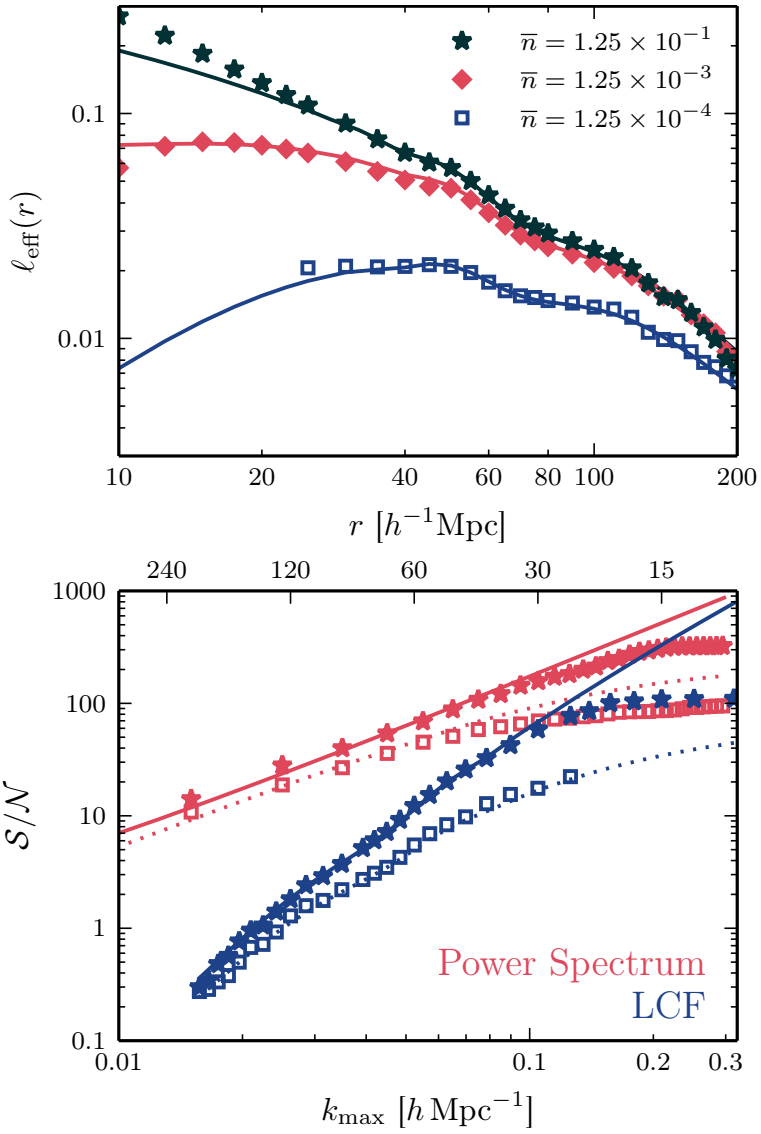


Fig. 5.4 *Upper panel:* Estimated LCF at redshift $z = 0.0$ for different particle densities (in units of $h^{-3}\,\mathrm{Mpc}^3$), corrected for additive shot noise terms. Stars mark the original measurements using all particles, diamonds (squares) derive from a subsample with 1 % (0.1 %) of the particles. Solid lines in the same colour correspond to the tree-level predictions. *Lower panel:* Measured cumulative signal-to-noise for power spectrum (red) and LCF (blue), compared to the approximation with Gaussian errors (solid and dotted lines). Symbols are the same as in the upper panel.

but coming from a subsample of particles that is randomly selected before the particles are smoothed onto the grid. This procedure does not correspond to a Poisson sampling of the underlying matter field, but it should serve as a close enough approximation thereof. In the upper panel of Fig. 5.4 diamonds (squares) show the measured LCF for 1 % (0.1 %) of the total number of particles, compared to the original measurements (stars), whereas the solid lines of matching color are the theoretical predictions in tree-level SPT based on the bispectrum term in Eq. (5.22). We clearly see how the measurements

get increasingly suppressed with decreasing number density and accurately follow the predictions, confirming our model.

5.4.6 Testing the signal-to-noise of $\hat{\ell}$

The lower panel of Fig. 5.4 contrasts the measured cumulative signal-to-noise with the idealized case of Gaussian errors presented in Sec. 5.3.2, where the symbols are the same as in the upper panel and colors distinguish between power spectrum and LCF. For both number densities we notice that the measured signal-to-noise traces the predicted one very well up to a k_{max} of $\sim 0.1\,h\,\mathrm{Mpc}^{-1}$. Beyond this scale the measured signal-to-noise quickly flattens out because higher-order corrections to the power spectrum and LCF covariance diminish the amount of available information. As can be seen in the plot, this effect can be quite severe: at the largest k_{max} the power spectrum signal-to-noise is approximately reduced by a factor of three, while the LCF signal-to-noise even suffers by a factor of seven.

5.4.7 Estimating the covariance matrix

Apart from the means, it is also instructive to consider the covariance matrix of the power spectrum and LCF estimators. To begin with, in Fig. 5.5 we compare the measured LCF variance in all 30 bins with the Gaussian approximation (dashed line) from Eq. (5.43) for the same three redshifts as in Fig. 5.3. The error bars were estimated via Jackknife resampling, meaning we first computed the variance σ_ℓ^2 from the full sample of $N_{\mathrm{real}} = 200$ realizations and subsequently from N_{real} different subsamples, each giving $(\sigma_\ell^{(i)})^2$, in which the i-th realization has been left out. The error on the variance is then computed as follows (Norberg et al., 2009):

$$\delta\sigma_\ell^2 = \sqrt{\frac{N_{\mathrm{real}}-1}{N_{\mathrm{real}}}\sum_{i=1}^{N_{\mathrm{real}}}\left[\left(\sigma_\ell^{(i)}\right)^2-\sigma_\ell^2\right]^2}\,. \tag{5.64}$$

From the figure we learn that on large scales higher-order variance terms are clearly negligible and the measured variance displays the expected r^3-scaling of the Gaussian approximation (Eq. 5.43). However, at a scale of $\sim 30\,h^{-1}\mathrm{Mpc}$ this agreement breaks down, the variance reaches a minimum and starts increasing with declining scales, which marks the onset of higher-order corrections. While the Gaussian variance is independent of redshift, higher-order terms are not and thus higher redshifts show smaller deviations from the Gaussian approximation. The scale at which the higher-order terms become important coincides with the scale where the measured signal-to-noise was observed to flatten out in Sec. 5.4.5 but also with the point at which tree-level SPT breaks down (cf. Fig. 5.3).

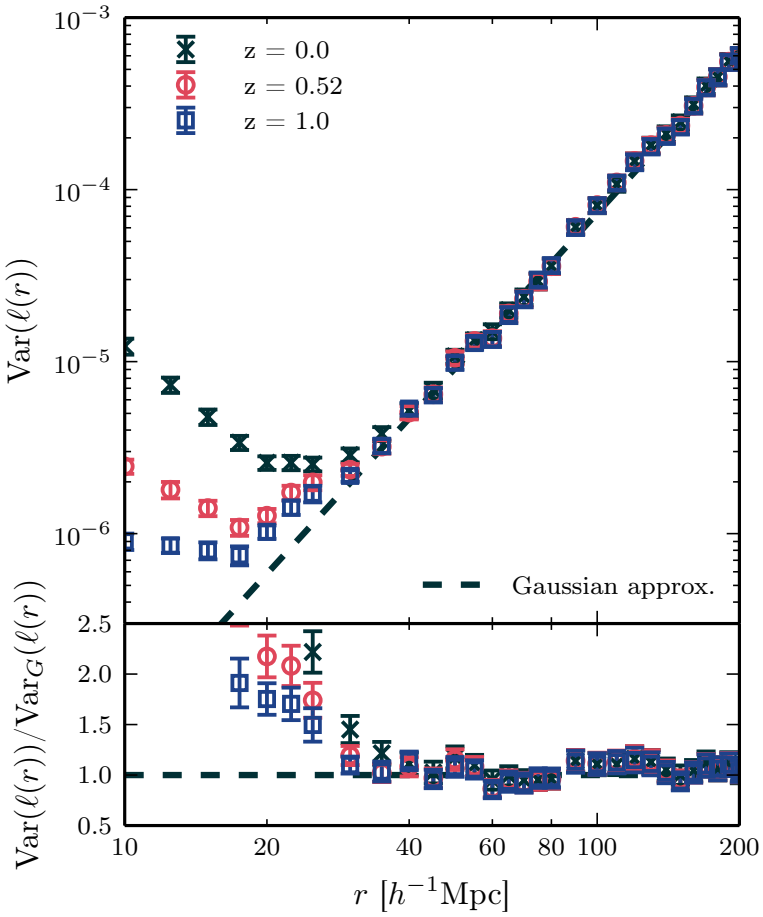


Fig. 5.5 Variance of the LCF estimator compared to the Gaussian approximation $\mathrm{Var}_G(\ell(r))$ from Eq. (5.43) for the same three redshifts as in Fig. 5.3. Error bars originate from a Jackknife resampling of the 200 realisations.

Figure 5.6 shows the full auto- and cross-correlation matrices for the power spectrum (bottom left panel) and LCF (top right panel), where the correlation coefficient r_{ij} is defined to be:

$$\mathsf{r}_{ij} = \frac{\langle \delta X_i \, \delta X_j \rangle}{\sqrt{\langle \delta X_i^2 \rangle \, \langle \delta X_i^2 \rangle}} \,. \tag{5.65}$$

The figure presents the measurements from simulations as all of the bins below the diagonal, whereas all bins above denote the theoretical prediction from the respective lowest-order contributions. In the case of the prediction for the cross-correlation (top left quadrant) this results from the bispectrum term in Eq. (5.51), computed at tree-level in SPT. Apart from some noise in the measurements we notice that both auto-correlation matrices are very well reproduced by their Gaussian approximations on large scales. On smaller scales, though, different bins become increasingly correlated with each other, which is underpredicted by the lowest-order contributions.

The measured cross-correlation matrix (bottom right quadrant) indicates that the power spectrum and LCF are largely but not entirely independent of each other. The

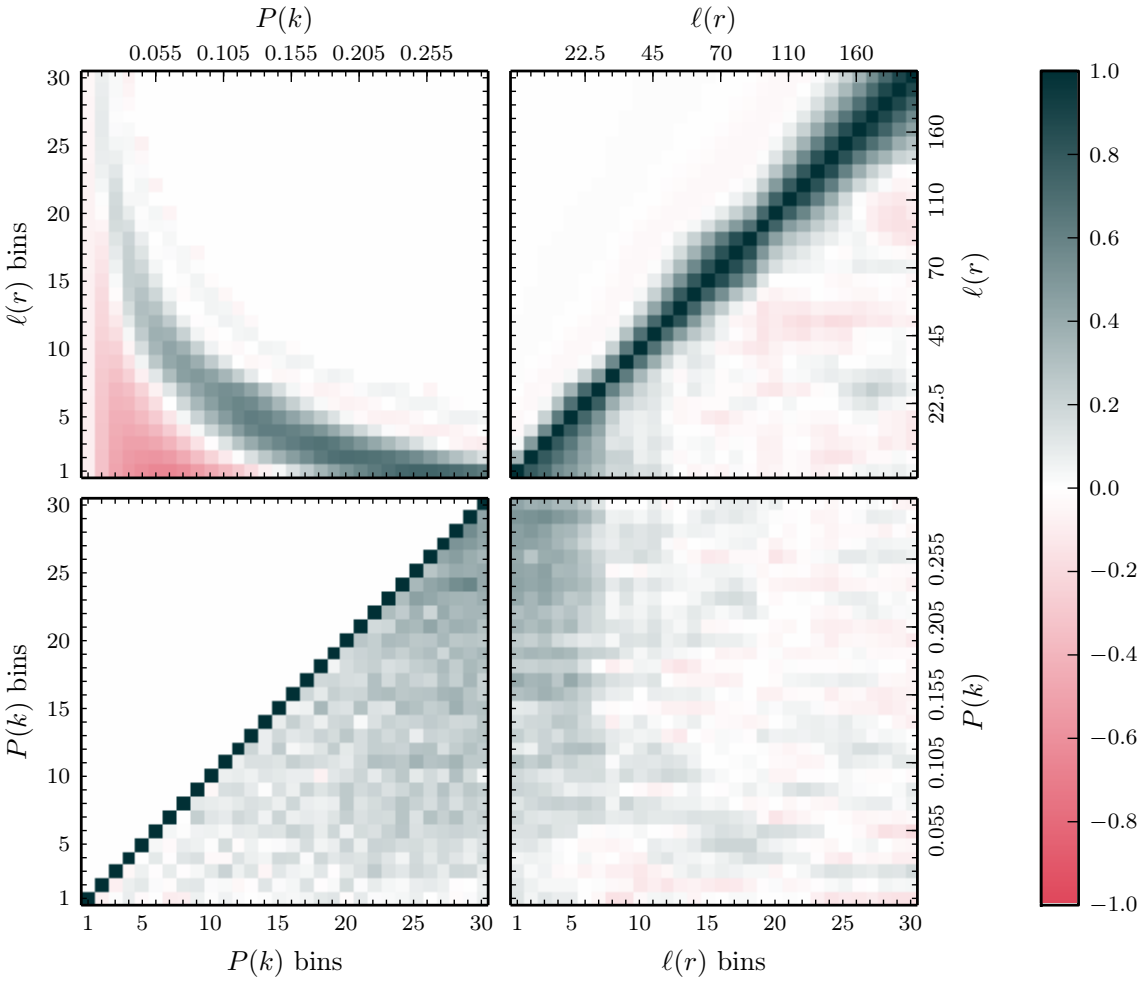


Fig. 5.6 Correlation matrices at redshift $z = 0$ with the auto-correlation of the LCF in the bottom left and the power spectrum in the top right panel. All bins below the diagonal derive from the measurements, all bins above are predictions based on the lowest-order contributions to either the auto- or cross-covariance and linear perturbation theory. Note that on power spectrum axes smaller scales (higher k) are to the right, whereas for the LCF these lie on the left.

small scale LCF bins seem to be reasonably correlated with most of the power spectrum bins. A qualitatively similar behaviour can be seen from the theoretical computation in the top left quadrant, which displays an arc with moderate correlations for power spectrum and LCF bins that are related by $k = \pi/r$. On large scales these correlations are of the order ~ 0.2 and are therefore only hardly identifiable in the measured data, but we do recognize positive correlations along the position of the arc. For small LCF scales and small power spectrum modes the cross-correlation is predicted to become negative, which is not seen in the data. However, in this regime we have to expect the breakdown of tree-level SPT as well as the influence of the next-order term in Eq. (5.51).

5.5 Detectability of the LCF in future surveys

5.5.1 Detectability of the LCF in galaxy survey data

Before we move on to discuss the LCF's sensitivity on various cosmological parameters, we consider the significance at which the LCF might be detected in a hypothetical galaxy survey. Taking the null hypothesis to be the absence of any signal, the χ^2 of a detection is simply given by

$$\chi^2_\ell = \sum_{i,j=1}^{N_\mathrm{bin}} \hat{\ell}_i \, \hat{\mathsf{C}}^{-1}_{ij} \, \hat{\ell}_j \,. \tag{5.66}$$

Using all 30 bins of our measurements we obtain $\chi^2_\ell \approx 10^4$, where we have accounted for the fact that the inverse of the estimated covariance matrix is not an unbiased estimate of the inverse and applied the Anderson–Hartlap factor (Hartlap et al., 2007), such that

$$\hat{\mathsf{C}}^{-1} = \frac{N_\mathrm{real} - N_\mathrm{bin} - 2}{N_\mathrm{real} - 1} \, \hat{\mathsf{C}}^{-1}_* \,, \tag{5.67}$$

where $\hat{\mathsf{C}}^{-1}_*$ is the algebraic inverse of the measured covariance matrix. The χ^2 above is the expected value for a measurement from a single simulation box, which has a volume of $V_\mathrm{box} = (1.5\,h^{-1}\mathrm{Gpc})^3$. As the errors scale inversely with volume, the χ^2_ℓ for a survey of volume V and an ideal box-like geometry is thus

$$\chi^2_\ell(V) = 10^4 \, \frac{V}{V_\mathrm{box}} \,. \tag{5.68}$$

A 5-σ detection from 30 data points corresponds to a $\chi^2 \sim 85$ and this could already be achieved by a survey with volume $V \approx 0.03\,h^{-3}\mathrm{Gpc}^3$.

5.5.2 Detectability of BAO features in the LCF

As was noted in previous plots (e.g. see Figs 5.1 and 5.3), the LCF signal displays slight wiggles. These are due to BAO imprinted in the matter distribution (for a review of the physics of BAO see Weinberg et al., 2013). It is therefore interesting to ask: what size does our idealized survey need to be in order to detect these features at a given confidence level? To answer this question we compute again the LCF in tree-level SPT, but this time for a featureless input power spectrum, obtained using the no-wiggle fitting formula of Eisenstein and Hu (1998). The relative difference $\left[\hat{\ell}(r) - \ell_\mathrm{nw}(r)\right] / \ell_\mathrm{nw}(r)$ between the measured and the no-wiggle LCF enables us to isolate the BAO features more clearly.

Figure 5.7 shows the results of this operation for our theoretical predictions from SPT (solid black line) and our measurements from simulations (crosses). The shape of this function can be understood as follows, the BAO signal in the two-point function has

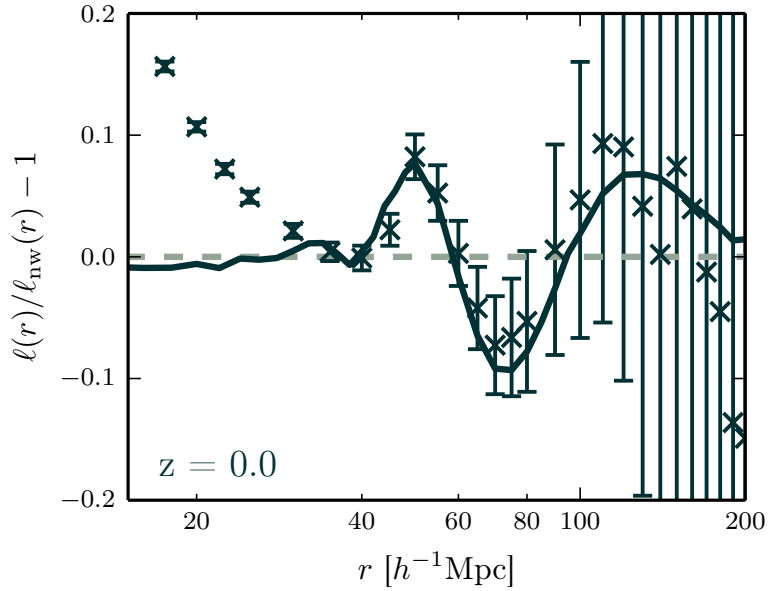


Fig. 5.7 Relative difference between measured and no-wiggle LCF, obtained from tree-level perturbation theory and a featureless power spectrum at redshift $z = 0.0$. The solid line represents the overall tree-level prediction. Error bars are scaled to match a survey of $V = 38\,h^{-3}\,\mathrm{Gpc}^3$.

a local maximum at roughly $r \sim 100\,h^{-1}$Mpc, which is an imprint of the sound horizon scale at recombination. Considering the LCF, this correlates three points along a line, each separated by distance r, and so there appears two values of r that would produce a resonance with the BAO scale: one when $r \sim 50\ h^{-1}$Mpc (i.e. the distance between points 3 and 1) and the second when $r \sim 100\ h^{-1}$Mpc (i.e the separation between points 2 and 1).

Analogously to the procedure of the previous subsection, let us take the standpoint that the no-wiggle LCF represents the null-hypothesis. Hence the χ^2 for detecting the BAO features can be written:

$$\chi^2_{\rm BAO} = \sum_{i,j=1}^{N_{\rm bin}} \left[\hat{\ell}_i - \ell_{\rm nw}(r_i)\right] \hat{\mathsf{C}}^{-1}_{ij} \left[\hat{\ell}_j - \ell_{\rm nw}(r_j)\right] . \tag{5.69}$$

Taking all 21 bins in the range from 40 to $200\,h^{-1}$Mpc we get $\chi^2_{\rm BAO} \approx 3.9$. Requiring a 3-$\sigma$ confidence level ($\chi^2 \approx 44$ for 21 data points), this translates into a minimal survey volume of $V \sim 38\,h^{-3}\mathrm{Gpc}^3$. This is within reach of upcoming galaxy surveys like DES, and certainly the Stage IV dark energy missions such as *Euclid* and *LSST*. For comparison, the BAO feature has already been detected in *BOSS* (comoving survey volume of $\sim 5\,h^{-3}\mathrm{Gpc}^3$) with a significance of about 10-σ from two-point statistics (Alam et al., 2017) and 4.5-σ using the ordinary three-point correlation function (Slepian et al., 2017b). While the BAO feature in the LCF might seem marginal compared to this, we reiterate that its information would be mostly complementary.

5.6 Cosmological Information

We now turn to address the question of the cosmological information content of the LCF, where the principal aim is to unveil which parameters or combination of parameters are best constrained by the LCF and also how it may help to tighten existing constraints obtained from the combination of the galaxy power spectrum and a Planck-like CMB measurement. Note that our main intention here is not to produce forecasts for a particular survey, but simply to provide a generic assessment of the possible relative gains from measuring the LCF. Thus in what follows we will make various simplifying assumptions. In paticular, we will neglect all effects that lead to anisotropies in the clustering of galaxies, such as redshift space distortions and the Alcock-Paczyński effect. Furthermore, we do not take any specifics of the galaxy population being surveyed into account and simply assume an ideal box-like geometry with a constant galaxy number density throughout. However, where possible we will try to add some validation for the choices that we make.

5.6.1 Formalism and assumptions

To begin we assume that the joint likelihood function for both the LCF and the power spectrum takes the form of a multi-variate Guassian:

$$\mathscr{L} = \frac{1}{\sqrt{(2\pi)^n |\mathbf{C}|}} \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \mathbf{C}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right], \tag{5.70}$$

where $\boldsymbol{x}$ is the vector containing the measured data, i.e $\boldsymbol{x}^T = \{\hat{P}_1, \ldots, \hat{P}_m, \hat{\ell}_1, \ldots, \hat{\ell}_n\}$ with mean $\boldsymbol{\mu} = \langle \boldsymbol{x} \rangle$, and $\mathbf{C}$ is the measured covariance matrix of dimension $(m+n) \times (m+n)$. Takahashi et al. (2009) have shown that the probability distribution of the power spectrum estimator is indeed very well approximated by a Gaussian distribution, over a broad range of scales.

For the case of the LCF there are no measurements in the literature to guide us, we therefore use our 200 realizations to determine the LCF probability distribution at four different scales, from 10 to 100 h^{-1}Mpc. Figure 5.8 shows the results form these set of measurements. The distribution is plotted as a function of $\delta\ell = \hat{\ell}_i - \langle \hat{\ell}_i \rangle$, normalized by the measured standard deviation, such that it should approach a Gaussian distribution with zero mean and unit variance (plotted as the thick, black line for reference). Albeit there is some scatter due to the small sample size, we do not observe any significant indication for a strong skewness or kurtosis and thus conclude that for our purposes here, the assumption of a Gaussian likelihood seems justified.

The parameter sensitivity of any statistic can be conveniently forecasted in the Fisher formalism. The Fisher matrix is obtained from the logarithm of the likelihood function

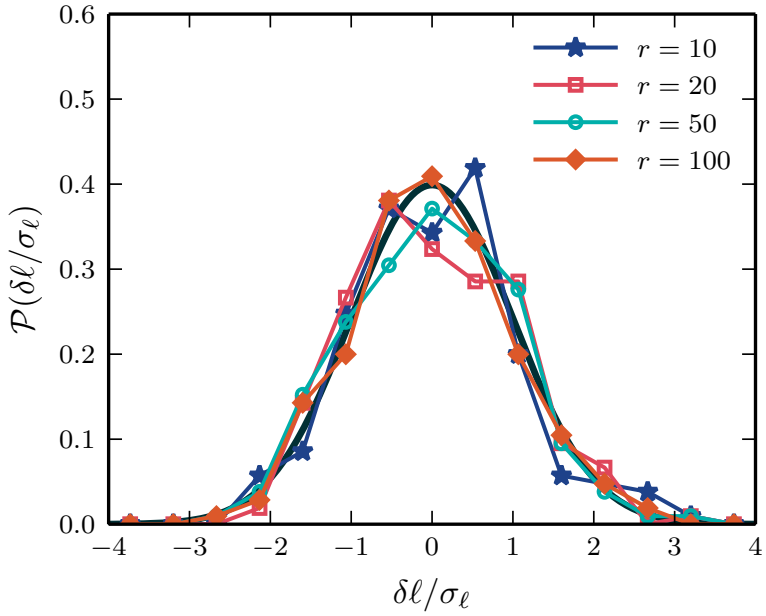


Fig. 5.8 Probability distribution of LCF estimator at redshift $z = 0.0$ as a function of the difference $\delta\ell = \hat{\ell}_i - \langle\hat{\ell}_i\rangle$, normalised by the standard deviation. Plotted are the results for various scales (given in units of h^{-1}Mpc) and a Gaussian with zero mean and unit variance (black line) for reference.

by taking second derivatives with respect to the parameters of interest θ_i,

$$\mathsf{F}_{ij} = -\left\langle \frac{\partial^2 \log \mathscr{L}}{\partial\theta_i \partial\theta_j} \right\rangle \Bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}, \tag{5.71}$$

where $\boldsymbol{\theta}_0$ denotes the set of fiducial parameter values. In particular, if the likelihood is Gaussian as in Eq. (5.70), it can be shown that the Fisher matrix takes the form (Tegmark et al., 1997)

$$\mathsf{F}_{ij} = \frac{1}{2}\mathrm{Tr}\left[\mathbf{C}^{-1}\frac{\partial \mathbf{C}}{\partial\theta_i}\mathbf{C}^{-1}\frac{\partial \mathbf{C}}{\partial\theta_j}\right] + \frac{\partial \boldsymbol{\mu}^t}{\partial\theta_i}\mathbf{C}^{-1}\frac{\partial \boldsymbol{\mu}}{\partial\theta_j}. \tag{5.72}$$

By computing Eq. (5.72) and then taking the inverse we get the minimal achievable error on a given parameter after marginalizing over all others,

$$\sigma(\theta_i) = \sqrt{\left(\mathsf{F}^{-1}\right)_{ii}}. \tag{5.73}$$

To facilitate the evaluation of Eq. (5.72) we can make a further approximation by neglecting the first term involving derivatives of the covariance matrix. For the case of the power spectrum one can argue that the second term on the right-hand-side of Eq. (5.72) scales directly with the number of Fourier modes, whereas the first term is independent and consequently is subdominant (Smith et al., 2014; Tegmark et al., 1997). For the case of the LCF this term vanishes identically at lowest order, since, as was shown earlier in

Table 5.1 Fiducial values of the cosmological parameters, along with the step sizes Δ each parameter has been varied in either direction in the simulations. The bias parameters are assumed to be $b_1 = 1$ and $b_2 = 0$.

Parameter	Ω_m	Ω_b	w_0	w_a	σ_8	n_s	h
Fiducial value	0.25	0.040	-1.0	0.0	0.8	1.00	0.70
Increment $\boldsymbol{\Delta}$	± 0.05	± 0.005	± 0.2	± 0.1	± 0.1	± 0.05	± 0.05

Eq. (5.38), the Gaussian part of the covariance is independent of cosmology and as was demonstrated in Fig. 5.5 the Gaussian part was shown to be a reasonable approximation for a wide range of scales.

5.6.2 Parameter sensitivity

In the following we consider a set of nine parameters,

$$\boldsymbol{\theta} = \{\Omega_m, \Omega_b, w_0, w_a, \sigma_8, n_s, h, b_1, b_2\}\,, \tag{5.74}$$

comprising the total matter and baryon densities Ω_m, Ω_b, the dark energy equation of state parameters w_0 and w_a, the amplitude of density fluctuations in spheres of $8\,h^{-1}$Mpc, the scalar spectral index n_s and the dimensionless Hubble rate h. Lastly, we also include the two bias parameters b_1 and b_2, using the local Lagrangian bias model introduced in Sec. 5.2.3. The fiducial values that we are adopting for each of the parameters are summarized in Table 5.1.

In order to compute the Fisher matrix, we now need to determine how the power spectrum and LCF respond to changes in these parameters, that is we need to evaluate the respective derivatives. To do that, we first generate modified linear power spectra where one of the parameters has been changed by a step up or down according to the values given in Table 5.1, while all others are kept at the fiducial value. Using these spectra we compute the LCF in the tree-level approximation of Eq. (5.12) and additionally all one-loop power spectra. That is necessary in order to include the bias parameter b_2 in the power spectrum Fisher matrix because at linear order it only depends on b_1. The corrections due to non-linear and non-local bias terms are summarized in App. 5.B. The derivatives are finally obtained by taking the central finite difference of the upward and downward steps, i.e.

$$\frac{\mathrm{d}X_i(\boldsymbol{\theta})}{\mathrm{d}\theta_\alpha} \approx \frac{X_i(\boldsymbol{\theta} + \Delta\theta_\alpha) - X_i(\boldsymbol{\theta} - \Delta\theta_\alpha)}{2\,\Delta\theta_\alpha}\,, \tag{5.75}$$

whereas for the two bias parameters we calculate the exact derivatives from Eqs. (5.16) and (5.102).

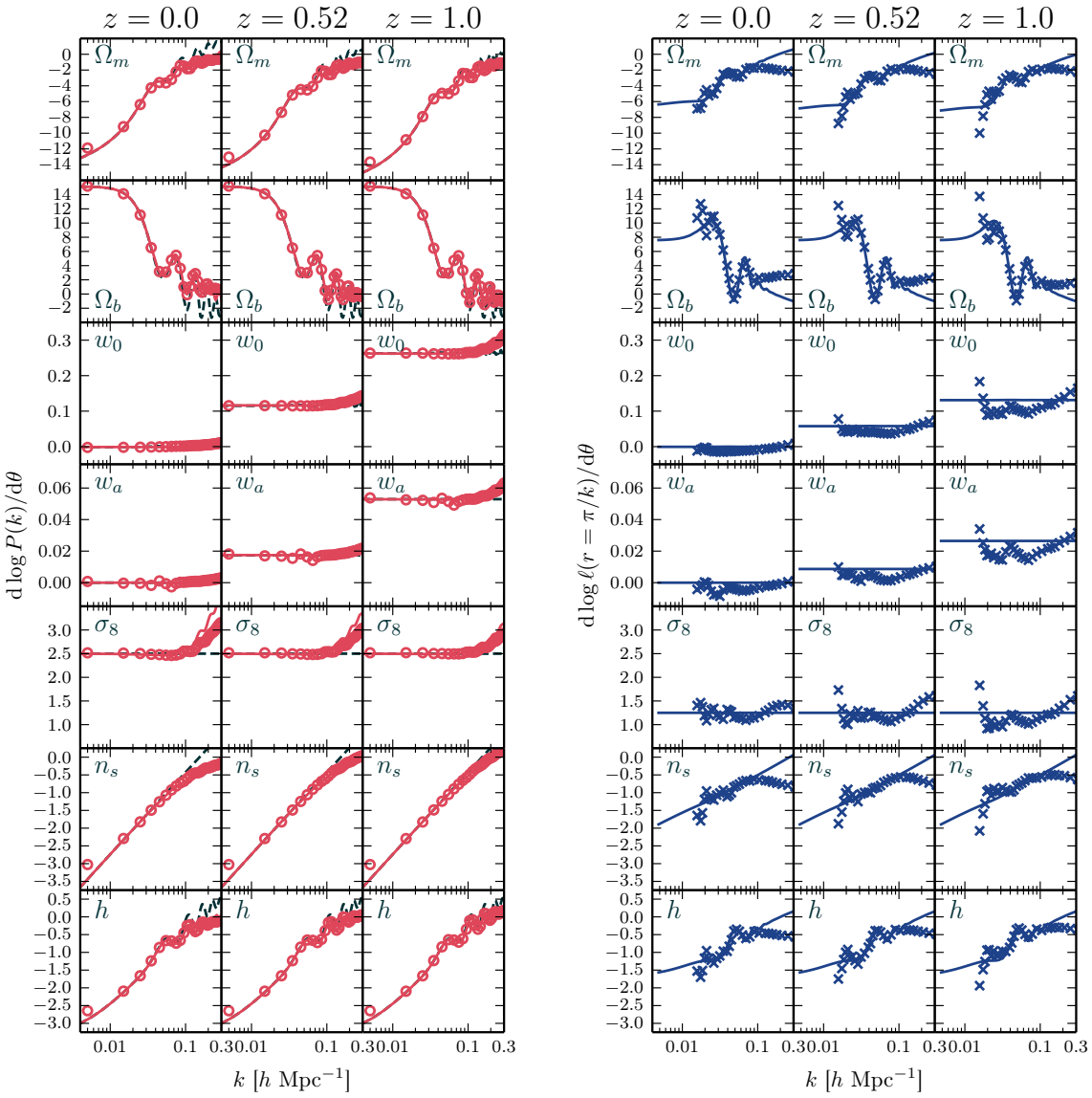


Fig. 5.9 Time evolution of the logarithmic derivatives of power spectrum (left) and LCF (right) with respect to various cosmological parameters. Data points represent direct measurements from the N-body simulations described in the text, while solid lines are the respective model predictions. For comparison, in the left panel we also show the linear power spectrum derivatives as the black, dashed lines.

To check the accuracy of these model predictions, we also measure the derivatives directly from a set of simulations whose cosmological parameters are varied in the same way as the ones given in Table 5.1 (originally performed in Smith et al., 2014). The specifics of the simulations are the same as the ones described in Sec. 5.4.1 and for each variation as well as the fiducial parameter set there are four realizations. To reduce the effect from sample variance, the phases of the initial Gaussian random field of each realization are matched to the corresponding one from the fiducial model. Derivatives are estimated as in Eq. (5.75) and averaged over the four realizations.

Figure 5.9 shows the comparison of our predictions for the logarithmic derivatives (depicted by solid lines) with the measurements from the N-body simulations (circles and crosses), and for the various cosmological models considered. The left panel shows the derivatives of the power spectrum and the right the LCF. In the left panel of Fig. 5.9 we see that the power spectrum derivatives are reasonably well captured by the one-loop model up to the maximal scale that is being considered, $k = 0.3\,h\,\mathrm{Mpc}^{-1}$. Note that for reference we also show the linear theory derivatives, indicated as the dashed lines. Considering the right panel of Fig. 5.9 we see that the tree-level predictions for the LCF are in reasonable agreement with the simulations up to $k = 0.1\,h\,\mathrm{Mpc}^{-1}$, beyond which the measured data display, in absolute terms, a larger derivative than the one predicted.

On comparing the power spectrum derivatives with those of the LCF, we find that the former usually dominate over the latter. For the three parameters w_0, w_a and σ_8, which mainly affect the amplitude of the power spectrum, the difference is a factor of ~ 2 over the scales considered. This matches our expectations, since as may be noted from inspecting Eq. (5.12) the LCF scales with the square root of the power spectrum amplitude. On the other hand, for the remaining four parameters we find that the power spectrum derivatives are only larger than the LCF ones on very large scales. On smaller scales the derivatives approach zero, signalling that the power spectrum does not provide much information on these parameters in the non-linear regime. The LCF model also predicts nearly vanishing derivatives on small scales but the fully non-linear measurements all saturate at some value, such that it might still possible to gain information. However, we note that mode coupling in the non-linear regime tends to erase memory of the initial conditions. If a parameter, such as Ω_b, is only relevant for determining the density at the initial time, it seems counter-intuitive that an observable would still be sensitive to this parameter at high k values. Further investigations are therefore required to discern whether the effect seen in the LCF is genuine or caused by systematics in the simulations and/or measurement process.

5.6.3 Forecasted parameter accuracy

To compute the Fisher matrix we assume that our idealistic survey consists of three independent redshift slices at $z = 0.0$, 0.52 and 1.0, each of a volume $V = 3.375\,(h^{-1}\mathrm{Gpc})^3$, so that the total Fisher matrix based on large-scale structure is given by

$$\mathbf{F}^{\mathrm{LSS}} = \mathbf{F}(z = 0.0) + \mathbf{F}(z = 0.52) + \mathbf{F}(z = 1.0)\,. \tag{5.76}$$

We take the theory predictions presented in Sec. 5.6.2 to model the parameter derivatives but use the fully non-linear covariance matrices estimated from the large suite of N-body simulations and correct the inverse for the Anderson–Hartlap factor (see Sec. 5.5). When considering combinations of power spectrum and LCF we use a bit more caution

when calculating the inverse as the two statistics have signals of widely differing orders of magnitudes. Consequently, the entries in the combined covariance matrix will equally vary by large amounts, making the inversion process subject to numerical errors. For that reason we first compute the correlation matrix $\mathbf{r}$, whose entries all lie in the interval $[-1, 1]$, and obtain its inverse via a singular value decomposition. The (uncorrected) inverse of the covariance matrix can then be written as (Smith et al., 2014)

$$\mathsf{C}^{-1}_{*,ij} = \frac{\mathsf{r}^{-1}_{*,ij}}{\sigma_i \sigma_j} \,. \tag{5.77}$$

Furthermore, we add the information coming from a CMB experiment like *Planck*, acting as priors for our parameter set. For that we initially compute the CMB Fisher matrix in a different parameter set that is more suitable for the CMB and then transform this matrix to match our chosen large-scale structure parameters (for more details, see App. A of Smith et al., 2014). We treat the CMB information as independent from the large-scale structure and hence the total Fisher matrix is finally given by

$$\mathbf{F}^{\rm tot} = \mathbf{F}^{\rm LSS} + \mathbf{F}^{\rm CMB} \,. \tag{5.78}$$

Figure 5.10 shows the 1σ likelihood contours derived from this Fisher matrix for various combinations of parameters, after marginalizing over all others, and a maximal mode $k_{\rm max} = 0.3\,h\,{\rm Mpc}^{-1}$ ($r_{\rm min} \sim 10\,h^{-1}{\rm Mpc}$). The error ellipses are constructed by inverting $\mathbf{F}^{\rm tot}$ and reducing it to a 2×2 submatrix of the desired parameters. This submatrix is inverted back again and we determine its eigenvalues and eigenvectors, which are used as input to plot the corresponding error ellipses. In each panel the red lines represent the case where $\mathbf{F}^{\rm LSS}$ is evaluated for the power spectrum alone, the blue dashed lines are for the LCF and the black ones the combination of both measures. Note that the CMB prior is always added.

The figure illustrates that there are some substantial gains over the power spectrum plus CMB alone to be made. In particular, the largest gains are obtained for the σ_8 parameter, which is mainly a proxy for the amplitude of fluctuations, and the two bias parameters b_1 and b_2. There is also a more modest improvement in the constraints on the matter density parameter $\Omega_{\rm m}$. However, for the other parameters $\{\Omega_b, n_s, h, w_0, w_a\}$ the gains are marginal. For the first three parameters, this is not too surprising since they are already well constrained by the CMB. The dark energy equation of state parameters do not display a significant improvement, either. Perhaps this owes to the fact that the LCF is only very weakly dependent on growth history and the nonlinear interaction kernel seems to be somewhat cosmology independent.

The above qualitative findings are shown more quantitatively in Tables 5.2 and 5.3, which summarizes the marginalized 1σ errors for all parameters and for two different

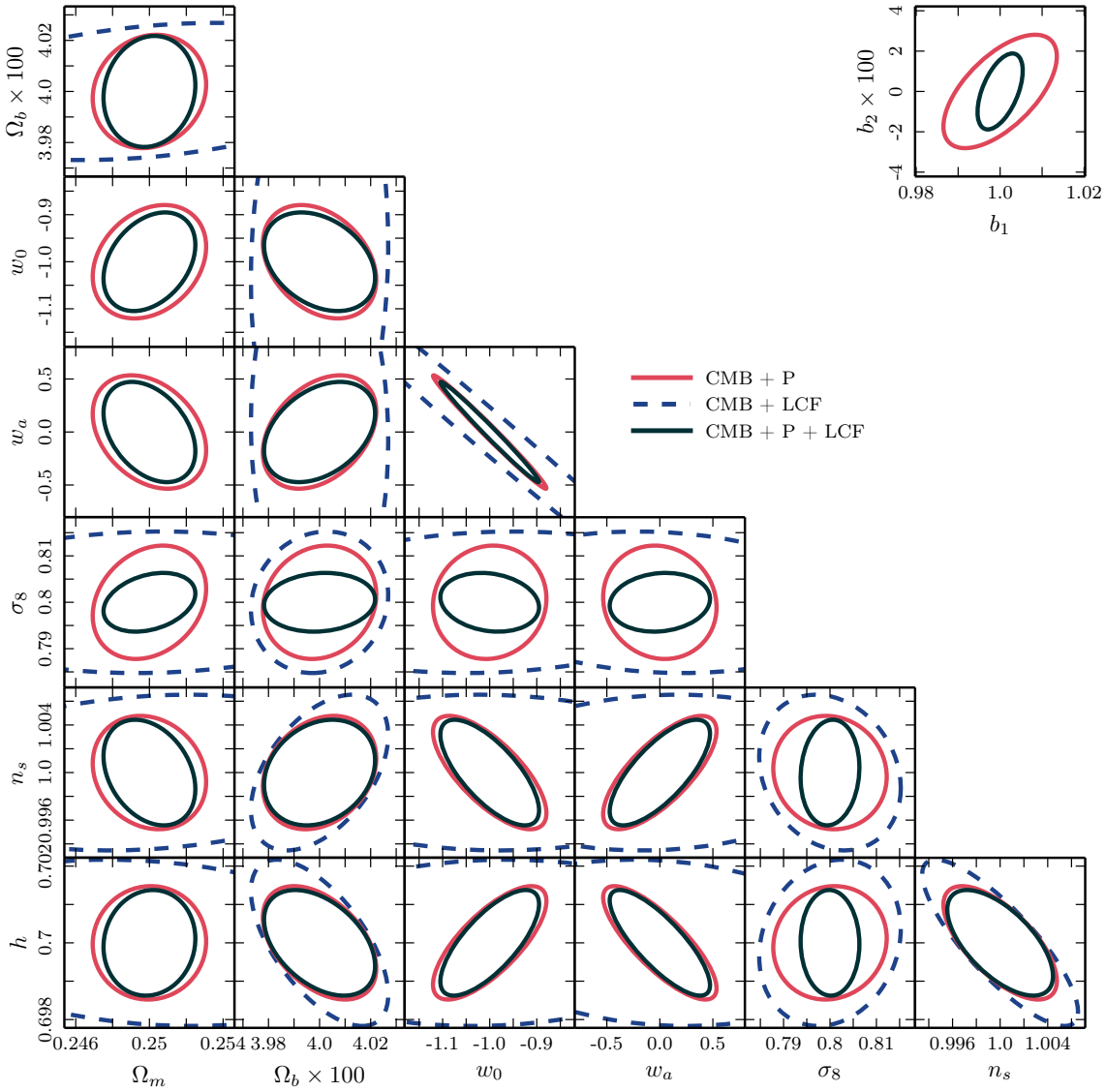


Fig. 5.10 Forecasted 1σ likelihood contours for various combinations of parameters, marginalised over all others. Power spectrum forecasts are represented by the red lines, the LCF by the blue dashed ones and their combination is shown in black. All forecasts use a cutoff scale $k_{\rm max} = 0.3\,h\,{\rm Mpc}^{-1}$ (corresponding to $r_{\rm min} \sim 10\,h^{-1}{\rm Mpc}$) and include information from a *Planck*-like CMB experiment.

cut-off scales: the one used for Fig. 5.10 as well as the more conservative choice $k_{\rm max} = 0.2\,h\,{\rm Mpc}^{-1}$. The combinations of power spectrum and LCF (including the CMB priors) are to be found in the sixth and second data column, respectively, where the percentages in parenthesis give the improvement compared to the power spectrum alone and we read off that the errors for σ_8 and b_1 decrease by 48 and even 60 % when LCF information is included. Interestingly, we obtain comparable improvement factors when the lower cut-off is being used, only b_2 displays a larger change from 13 to 33 %, indicating that the information coming from smaller LCF scales are particularly helpful in constraining non-linear bias. However, we note that the improvement factors are slightly underestimated

Table 5.2 Marginalised 1σ errors for the power spectrum and a combination of power spectrum and LCF (CMB priors are included in all cases) using the cutoff scale $k_{\max} = 0.2\,h\,\mathrm{Mpc}^{-1}$ ($r_{\min} \sim 16\,h^{-1}\mathrm{Mpc}$). The last two columns correspond to the case where the two bias parameters have been fixed to their fiducial values. The percentages in the parenthesis indicate the improvement over the respective power spectrum results.

	P	P + LCF	P fixed bias	P + LCF fixed bias
$\Delta\Omega_m$	0.0022	0.0019 (14 %)	0.0016	0.0015 (4 %)
$\Delta\Omega_b$	0.000153	0.000151 (1 %)	0.000151	0.000150 (1 %)
Δw_0	0.094	0.084 (11 %)	0.088	0.083 (6 %)
Δw_a	0.401	0.370 (8 %)	0.388	0.369 (5 %)
$\Delta\sigma_8$	0.0096	0.0060 (38 %)	0.0012	0.0011 (9 %)
Δn_s	0.0035	0.0033 (4 %)	0.0034	0.0033 (1 %)
Δh	0.00109	0.00106 (3 %)	0.00108	0.00106 (2 %)
Δb_1	0.012	0.005 (60 %)	–	–
Δb_2	0.023	0.020 (13 %)	–	–

Table 5.3 Same as Table 5.2 but for the cut-off scale $k_{\max} = 0.3\,h\,\mathrm{Mpc}^{-1}$.

	P	P + LCF	P fixed bias	P + LCF fixed bias
$\Delta\Omega_m$	0.0020	0.0016 (19 %)	0.00122	0.00116 (5 %)
$\Delta\Omega_b$	0.000146	0.000143 (2 %)	0.000145	0.000142 (2 %)
Δw_0	0.079	0.069 (13 %)	0.078	0.069 (12 %)
Δw_a	0.352	0.311 (12 %)	0.347	0.311 (10 %)
$\Delta\sigma_8$	0.0080	0.0042 (48 %)	0.0010	0.0008 (21 %)
Δn_s	0.0031	0.0029 (7 %)	0.0030	0.0029 (4 %)
Δh	0.00097	0.00091 (7 %)	0.00095	0.00090 (5 %)
Δb_1	0.009	0.004 (60 %)	–	–
Δb_2	0.018	0.012 (33 %)	–	–

because as we have seen in Sec. 5.6.2, our LCF model somewhat underpredicts the parameter sensitivity in the non-linear regime.

We also consider the case where we assume that the bias parameters are known and fixed to their fiducial values, meaning that we simply strike out all the corresponding rows and columns in $\mathbf{F}^{\mathrm{tot}}$. The resulting errors are given in the last two columns for each cut-off scale and we now observe a reduction of all improvement factors, which is particularly evident for Ω_m and σ_8. That implies that for these parameters the gain from the LCF is mainly due to a better constraint of galaxy bias. We also see that the inclusion of a larger amount of the small scale modes now brings about slightly more significant improvements.

To further investigate where the main constraining power of the LCF is coming from, we analyze a subset of our parameters that just comprises σ_8 and the bias parameters.

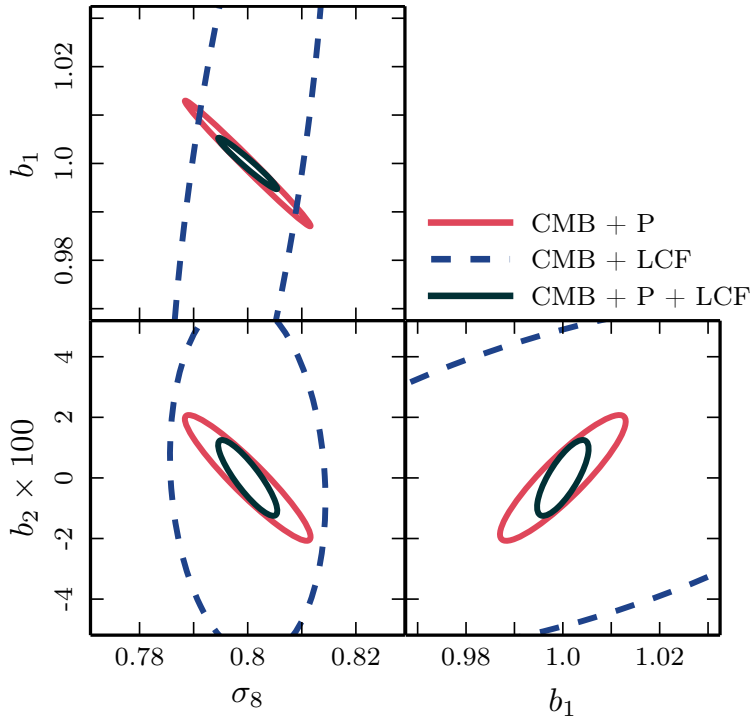


Fig. 5.11 Forecasted 1σ likelihood contours as in Fig. 5.10 but for a subset of parameters including σ_8, b_1 and b_2. The cut-off scale is chosen to be $k_{\max} = 0.3\,h\,\mathrm{Mpc}^{-1}$.

The resulting likelihood contours of an analogue Fisher matrix computation are shown in Fig. 5.11 and they clarify why we obtain the comparably good improvements noted above. Even though the LCF cannot put a tight constraint on b_1, it does produce small error bars for σ_8 (cf. $\Delta\sigma_8$ in the last column of Tables 5.2 and 5.3). As the degeneracy direction in the σ_8-b_1 plane (see top left panel in Fig. 5.11) for the LCF is fundamentally different from that of the power spectrum, a combination of both measures lead to good constraints, which in turn carry over to b_2. This behaviour is reasonable because as was shown in Sec. 5.2.3 the LCF is independent of linear bias at lowest order (but not of σ_8) and thus breaks the degeneracy between both parameters when combined with e.g. the power spectrum.

5.7 Discussion and Conclusions

In this chapter, we have studied the ability of the LCF to constrain the cosmological model in combination with power spectra measurements of the large-scale galaxy and CMB fluctuations.

In order to achieve this it was necessary to extend the LCF from describing matter fluctuations to those of galaxy fluctuations. In Sec. 5.2 we did this by computing the LCF in the Lagrangian biasing scheme. While the LCF is independent of bias in the regime where the relation between the galaxy and dark matter over-densities is linear, we have seen that non-linearity and non-locality introduce additional terms. However,

if bias is assumed to be local in Lagrangian space (this approach was adopted in recent bispectrum measurements from *BOSS* by Gil-Marín et al., 2017), the LCF still breaks the degeneracy between the amplitude of density fluctuations and the two remaining bias parameters. In comparison, this is not possible if one considers the bispectrum alone.

We also determined the effect of shot noise on the LCF, finding that after additive contributions are removed the signal becomes increasingly suppressed with decreasing number densities. Unlike more conventional clustering measures, the LCF cannot be completely cleaned from shot noise and the galaxy number density must hence be incorporated into the modelling.

In Sec. 5.3 we provided the first ever derivation of the LCF auto-covariance and its cross-covariance with the power spectrum. We noted that there was a structural similarity for the joint covariance we computed and that associated with the joint covariance between the power spectrum and the bispectrum. More importantly, though, we were able to prove that, in the Gaussian limit, the lowest-order contribution to the LCF covariance was independent of cosmological parameters. For that reason it is not subject to non-linear evolution and can therefore be predicted exactly. This property might prove to be advantageous compared to the bispectrum, as it might allow us to produce accurate covariance matrices from relatively cheap small-scale N-body simulations, matched with the analytic results in the Gaussian limit. Of course there are additional higher-order corrections which involve the trispectrum, quadratic powers of the bispectrum, and the 6-point spectrum of the phase field that could complicate this possibility.

In Sec. 5.4 we confronted our analytic results with measurements from a large ensemble of N-body simulations, comprising a total combined volume of $675\,h^{-3}\,\mathrm{Gpc}^3$. The simulations enabled us to produce the first ever measurement of the LCF at the BAO scale, with enough volume, to unambiguously detect the acoustic oscillation features. Through the simulation to simulation variance, we were also able to produce the first ever measurement of the LCF covariance matrix and its cross-covariance with the power spectrum. We found that the Gaussian approximation of the covariance holds down to scales $\sim 40\,h^{-1}\mathrm{Mpc}$, slightly above the scale where the tree-level SPT result for the LCF breaks down. As expected, the power spectrum and LCF are mostly uncorrelated but we do detect moderate correlations of the order $\lesssim 0.4$ for small LCF bins as well as bins of equivalent scales, i.e. those, which are related via $k \approx \pi/r$.

In addition, we also computed the signal-to-noise ratio of the LCF, from which we discovered that the power spectrum dominates over the LCF up to $k_{\mathrm{max}} = 0.3\,h\,\mathrm{Mpc}^{-1}$ ($\sim 10\,h^{-1}\mathrm{Mpc}$). However, the LCF signal-to-noise shows a stronger increase with decreasing scale. This suggests that there is potentially more information to be gained in the LCF than the power spectrum by pushing to smaller scales.

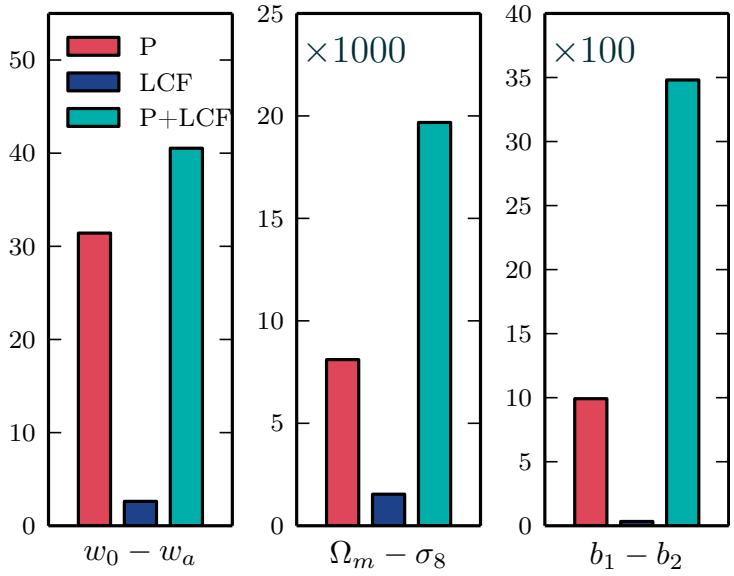


Fig. 5.12 Figure of merit for three different parameter combinations. The bar on the left in each panel is the power spectrum, the middle bar marks the LCF and the last one the combination of both. All three include the CMB priors.

In Sec. 5.5 we turned to assessing the detectability of the LCF in galaxy surveys. We found that the LCF could be measured at $> 5\sigma$ significance in a survey of volume $V > 0.03\,h^{-3}\mathrm{Gpc}^3$ – this paves the way for a very high significance measurement in modern surveys like the SDSS main sample and *BOSS*. On the other hand, we found that in order to detect the BAO signature one would require a survey that spans a volume of roughly $V \sim 40\,h^{-3}\mathrm{Gpc}^3$ and hence this would only likely be possible with future Stage IV missions like *Euclid* and *LSST*.

In Sec. 5.6 we explored the main question of this chapter, which was the cosmological information content of the LCF. We considered a large-scale structure survey of total volume $\sim 10\,h^{-3}\,\mathrm{Gpc}^3$, up to $z \sim 1$ – thus comparable with Stage-III like spectroscopic missions (Albrecht et al., 2006). We found that when estimates of the LCF are combined with estimates of the galaxy power spectrum and a Planck-like CMB experiment, significant improvements may be found in constraints on the amplitude of density fluctuations σ_8 (roughly a factor of ~2) and to more modest improvements in the matter density parameter Ω_m (~ 20%). In addition, one is able to significantly improve constraints on the nonlinear galaxy bias parameters b_1 and b_2 (factor ~ 2). Expressed in terms of the figure of merit (see Fig. 5.12) we obtain improvement factors of ~ 3.5 and ~ 2.4 for the parameter combinations b_1–b_2 and Ω_m–σ_8, respectively[1]. On

[1]The figure of merit is defined as the inverse area enclosed by the 2σ error ellipse for any combination of two parameters (Albrecht et al., 2006), i.e.

$$\mathrm{FOM}(\theta_i, \theta_j) \equiv \frac{1}{\pi\sqrt{6.17\,\mathrm{Det}\left[\mathbf{F}^{-1}(\theta_i, \theta_j)\right]}},$$

where $\mathbf{F}(\theta_i, \theta_j)$ is the 2×2 Fisher matrix of the parameters θ_i and θ_j.

the other-hand we find no significant improvement to be gained in the traditional dark energy figure of merit (it changes by a factor of ~ 1.3).

As we have only used 30 bins for our forecasts, these results suggest that the LCF provides an efficient compression of higher-order information. This could prove to be advantageous over the bispectrum, as it simplifies the task of generating accurate covariance matrices. Clearly, a more detailed study comparing various higher-order statistics is necessary to make a more definitive statement and will be presented in Chap. 6. Besides, by definition the LCF probes a very particular configuration of three points, so it would be interesting to explore whether there are more optimal shapes for constraining parameters.

Finally, to facilitate the Fisher analysis in this chapter, we have made a number of simplifications. In particular, we have neglected effects from redshift space distortions; the imprint of the finite survey geometry and survey mask on the phase field. Both of these will need to be explored in future work to arrive at more realistic forecasts and methodology for survey analysis. Accounting for redshift space distortions should provide additional information. Moreover, owing to the fact that in a power spectrum analysis the growth rate of structure $f(\Omega)$ is strongly degenerate with σ_8 and b_1, we expect the LCF will also prove to be effective in breaking these degeneracies and so *should* provide improvements in the dark energy figure of merit. Furthermore, we have assumed throughout that the primordial density field was Gaussian. If this was not the case, the LCF would acquire a further contribution, whose amplitude $f_{\rm NL}$ would be naturally degenerate with b_1, which means it can be potentially well constrained by a combination of power spectrum and LCF.

Appendices

Appendix 5.A The Joint Probability Density Function of Fourier Modes

In this appendix we present results from Matsubara (2003, 2007), which will be used for the derivation of the LCF covariance in Sec. 5.3.

Given a density field enclosed in a box of volume V with discrete Fourier modes $\delta_{\boldsymbol{k}}$, all of its statistical properties are encoded in the probability density function (PDF) $\mathscr{P}$. Normalising the density modes by volume and their power spectrum we get the dimensionless variables $\alpha_{\boldsymbol{k}} \equiv \delta_{\boldsymbol{k}}/\sqrt{V\,P(k)}$, in terms of which the PDF can be written as

$$\mathscr{P}[\alpha_{\boldsymbol{k}}] = \exp\left[\sum_{N=3}^{\infty} \frac{(-1)^N}{N!} \sum_{\boldsymbol{k}_1} \cdots \sum_{\boldsymbol{k}_N} \langle \alpha_{\boldsymbol{k}_1} \cdots \alpha_{\boldsymbol{k}_N} \rangle_c \frac{\partial}{\partial \alpha_{\boldsymbol{k}_1}} \cdots \frac{\partial}{\partial \alpha_{\boldsymbol{k}_N}}\right] \mathscr{P}_G[\alpha_{\boldsymbol{k}}]\,, \tag{5.79}$$

where

$$\mathscr{P}_G[\alpha_{\boldsymbol{k}}] = \frac{1}{2\pi} \exp\left[-\frac{1}{2}\sum_{\boldsymbol{k}} \alpha_{\boldsymbol{k}}\alpha_{-\boldsymbol{k}}\right] \tag{5.80}$$

denotes the Gaussian PDF. The $\langle \alpha_{\boldsymbol{k}_1} \cdots \alpha_{\boldsymbol{k}_N} \rangle_c$ refer to the N-th order cumulants,

$$\langle \alpha_{\boldsymbol{k}_1} \cdots \alpha_{\boldsymbol{k}_N} \rangle_c = p^{(N)}(\boldsymbol{k}_1, \ldots \boldsymbol{k}_N)\, \delta^K_{\boldsymbol{k}_1 + \cdots + \boldsymbol{k}_N}\,, \tag{5.81}$$

which are related to normalized (and dimensionless) versions of the ordinary N-th order spectra $P^{(N)}$, defined as

$$p^{(N)}(\boldsymbol{k}_1, \ldots \boldsymbol{k}_N) \equiv V^{1-\frac{N}{2}} \frac{P^{(N)}(\boldsymbol{k}_1, \ldots \boldsymbol{k}_N)}{\sqrt{P(\boldsymbol{k}_1) \cdots P(\boldsymbol{k}_N)}}\,. \tag{5.82}$$

For mildly non-Gaussian fields we can expand the exponential in Eq. (5.79),

$$\mathscr{P}[\alpha_{\boldsymbol{k}}] = \left[1 + \sum_{n=1}^{\infty} \mathscr{Q}^{(n)}\right] \mathscr{P}_G[\alpha_{\boldsymbol{k}}]\,, \tag{5.83}$$

where the terms $\mathcal{Q}^{(n)}$ contain spectra of various orders and can be expressed as follows

$$\mathcal{Q}^{(n)} = \sum_{m=1}^{\infty} \frac{1}{m!} \sum_{\substack{n_1,\dots,n_m \geq 1 \\ n_1+\cdots+n_m = n}} \frac{1}{(n_1+2)!\cdots(n_m+2)!} \sum_{\boldsymbol{k}_1^{(1)},\dots,\boldsymbol{k}_{n_1+2}^{(1)}} \cdots \sum_{\boldsymbol{k}_1^{(m)},\dots,\boldsymbol{k}_{n_m+2}^{(m)}} \times p^{(n_1+2)} \cdots p^{(n_m+2)} H_{\boldsymbol{k}_1^{(1)}\cdots\boldsymbol{k}_{n_1+2}^{(1)}\cdots\boldsymbol{k}_1^{(m)}\cdots\boldsymbol{k}_{n_m+2}^{(m)}}, \tag{5.84}$$

with H standing for a generalization of the Hermite polynomials,

$$H_{\boldsymbol{k}_1\cdots\boldsymbol{k}_n} = \frac{(-1)^n}{\mathscr{P}_G[\alpha_{\boldsymbol{k}}]} \frac{\partial}{\partial \alpha_{\boldsymbol{k}_1}} \cdots \frac{\partial}{\partial \alpha_{\boldsymbol{k}_N}} \mathscr{P}_G[\alpha_{\boldsymbol{k}}]\,. \tag{5.85}$$

In our Universe structure formation roughly follows the hierarchical model, meaning we have that $P^{(N)} \sim \mathcal{O}\left[P(k)^{N-1}\right]$, such that any given term in the series expansion above is of the order

$$\mathcal{Q}^{(n)} \sim \varepsilon^{N-2}, \qquad \varepsilon \equiv \sqrt{\frac{P(k)}{V}}. \tag{5.86}$$

It follows that the expansion is only meaningful as long as this parameter ε remains small. Furthermore, this allows us to conveniently keep track of the order of each term by counting powers of $1/\sqrt{V}$.

As we are going to evaluate ensemble averages comprising amplitudes and/or phases of Fourier modes, it is useful to split our variables accordingly and write them as $\alpha_{\boldsymbol{k}} = A_{\boldsymbol{k}}\,\mathrm{e}^{i\theta_{\boldsymbol{k}}}$. However, due to the reality constraint, $\alpha_{\boldsymbol{k}}$ and $\alpha^*_{\boldsymbol{k}} = \alpha_{-\boldsymbol{k}}$ are not entirely independent from each other, which is why we restrict all summations over wavevectors to the upper half sphere (uhs), defined by $k_z \geq 0$. In this subspace the probability for a set of modes to take values within an infinitesimal interval is thus given by

$$\mathscr{P}(\{\alpha_{\boldsymbol{k}}, \alpha^*_{\boldsymbol{k}}\}) \prod_{\boldsymbol{k}\in\text{uhs}} \mathrm{d}\alpha_{\boldsymbol{k}}\mathrm{d}\alpha^*_{\boldsymbol{k}} = \mathscr{P}(\{A_{\boldsymbol{k}}, \theta_{\boldsymbol{k}}\}) \prod_{\boldsymbol{k}\in\text{uhs}} 2A_{\boldsymbol{k}}\,\mathrm{d}A_{\boldsymbol{k}}\,\mathrm{d}\theta_{\boldsymbol{k}}\,, \tag{5.87}$$

where the factor $2A_{\boldsymbol{k}}$ comes from the Jacobian of the transformation. Expressing the $\mathcal{Q}^{(n)}$ in terms of $A_{\boldsymbol{k}}$ and $\theta_{\boldsymbol{k}}$, all resulting terms can be rearranged to display a similar structure:

$$\sum_{\substack{\boldsymbol{k}_1,\dots\in\text{uhs} \\ \boldsymbol{k}_i \neq \boldsymbol{k}_j}} A_{\boldsymbol{k}_1} A_{\boldsymbol{k}_2} \cdots \cos\left(\theta_{\boldsymbol{k}_1} \pm \theta_{\boldsymbol{k}_2} \pm \cdots\right) p^{(n_1)}(\boldsymbol{k}_1,\dots,\boldsymbol{k}_{n_1})\, p^{(n_2)}(\boldsymbol{k}_{n_1+1},\dots,\boldsymbol{k}_{n_1+n_2}) \cdots, \tag{5.88}$$

and it is important to note that when integrating over the phases, we always get a vanishing result unless the cosine-term is cancelled by some means. That is only possible if we correlate a number of phase factors, which exactly matches the number of phases appearing in Eq. (5.88). Since all other terms in Eq. (5.84) give no contribution, this drastically simplifies our task of computing any particular phase correlator. From this observation also follows that any even (odd) phase correlator can only get contributions

from even (odd) N-th order spectra. We will now consider some special cases, which occur in the main part of this work.

5.A.1 Four-point phase correlator

We only take into account terms of the order $1/V$, which corresponds to $n = 2$ in the PDF expansion of Eq. (5.83). After splitting each summation into separate sums over mutually different modes, we obtain the following two terms with four different phase factors,

$$\mathscr{Q}_1^{(2)} = \frac{1}{3} \sum_{\substack{\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3 \\ \boldsymbol{k}_i \neq \boldsymbol{k}_j}}^{\text{uhs}} A_{\boldsymbol{k}_1} A_{\boldsymbol{k}_2} A_{\boldsymbol{k}_3} A_{\boldsymbol{k}_{123}} \cos\left(\theta_{\boldsymbol{k}_1} + \theta_{\boldsymbol{k}_2} + \theta_{\boldsymbol{k}_3} - \theta_{\boldsymbol{k}_{123}}\right) p^{(4)}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, -\boldsymbol{k}_{123}), \tag{5.89}$$

$$\mathscr{Q}_2^{(2)} = \sum_{\substack{\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3 \\ \boldsymbol{k}_i \neq \boldsymbol{k}_j}}^{\text{uhs}} A_{\boldsymbol{k}_1} A_{\boldsymbol{k}_2} A_{\boldsymbol{k}_3} A_{\boldsymbol{k}_{123}} \cos\left(\theta_{\boldsymbol{k}_1} + \theta_{\boldsymbol{k}_2} - \theta_{\boldsymbol{k}_3} - \theta_{\boldsymbol{k}_{123}}\right) p^{(4)}(\boldsymbol{k}_1, \boldsymbol{k}_2, -\boldsymbol{k}_3, -\boldsymbol{k}_{123}), \tag{5.90}$$

where $\boldsymbol{k}_{123} \equiv \boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3$. We can trivially integrate over all amplitudes $A_{\boldsymbol{k}}$ and by making use of the identity

$$\mathscr{I}(n) = \int_0^\infty A^n 2A \mathrm{e}^{-A^2} \mathrm{d}A = \Gamma\left(1 + \frac{n}{2}\right), \tag{5.91}$$

we see that both expressions above acquire a factor of $\mathscr{I}(1)^4 = (\sqrt{\pi}/2)^4$ (note that $\mathscr{I}(0) = 1$). Let us now consider the correlator $\langle \epsilon_{\boldsymbol{q}_1} \epsilon_{\boldsymbol{q}_2} \epsilon_{\boldsymbol{q}_3} \epsilon_{-\boldsymbol{q}_{123}} \rangle$ with $\boldsymbol{q}_1$, $\boldsymbol{q}_2$, $\boldsymbol{q}_3$ all in the upper half sphere, such that a contribution from Eq. (5.89) looks as follows:

$$\begin{aligned} \propto & \sum_{\substack{\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3 \\ \boldsymbol{k}_i \neq \boldsymbol{k}_j}}^{\text{uhs}} p^{(4)}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, -\boldsymbol{k}_{123}) \int \frac{\mathrm{d}\theta_{\boldsymbol{q}_1}}{2\pi} \mathrm{e}^{i\theta_{\boldsymbol{q}_1}} \int \frac{\mathrm{d}\theta_{\boldsymbol{q}_2}}{2\pi} \mathrm{e}^{i\theta_{\boldsymbol{q}_2}} \int \frac{\mathrm{d}\theta_{\boldsymbol{q}_3}}{2\pi} \mathrm{e}^{i\theta_{\boldsymbol{q}_3}} \int \frac{\mathrm{d}\theta_{\boldsymbol{q}_{123}}}{2\pi} \mathrm{e}^{-i\theta_{\boldsymbol{q}_{123}}} \\ & \times \prod_{\boldsymbol{p} \neq \boldsymbol{q}_i}^{\text{uhs}} \int \frac{\mathrm{d}\theta_{\boldsymbol{p}}}{2\pi} \cos\left(\theta_{\boldsymbol{k}_1} + \theta_{\boldsymbol{k}_2} + \theta_{\boldsymbol{k}_3} - \theta_{\boldsymbol{k}_{123}}\right). \end{aligned} \tag{5.92}$$

As was already mentioned above, unless all phase factors are cancelled, the whole expression will evaluate to zero. Non-zero contributions therefore stem from cases where each $\boldsymbol{q}$-mode equals one of the $\boldsymbol{k}$-modes, giving in total $3! = 6$ different permutations. Eq. (5.90) does not add to this exemplary configuration and the full result is hence

$$\langle \epsilon_{\boldsymbol{q}_1} \epsilon_{\boldsymbol{q}_2} \epsilon_{\boldsymbol{q}_3} \epsilon_{-\boldsymbol{q}_{123}} \rangle = \left(\frac{\sqrt{\pi}}{2}\right)^4 p^{(4)}(\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3, -\boldsymbol{q}_{123}). \tag{5.93}$$

It can be checked that this holds true for all possible configurations of the $\boldsymbol{q}$-modes.

5.A.2 Six-point phase correlator

The lowest-order contributions to the connected six-point phase correlator come from terms with $n = 4$ in the PDF expansion. In this case we find three terms with six different phase factors,

$$\begin{aligned}\mathcal{Q}_1^{(4)} = \frac{1}{60} \sum_{\substack{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_5 \\ \boldsymbol{k}_i \neq \boldsymbol{k}_j}}^{\text{uhs}} & A_{\boldsymbol{k}_1} \cdots A_{\boldsymbol{k}_5} A_{\boldsymbol{k}_{12345}} \cos\left(\theta_{\boldsymbol{k}_1} + \theta_{\boldsymbol{k}_2} + \theta_{\boldsymbol{k}_3} + \theta_{\boldsymbol{k}_4} + \theta_{\boldsymbol{k}_5} - \theta_{\boldsymbol{k}_{12345}}\right) \\ & \times p^{(6)}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \boldsymbol{k}_4, \boldsymbol{k}_5, -\boldsymbol{k}_{12345}),\end{aligned} \tag{5.94}$$

$$\begin{aligned}\mathcal{Q}_2^{(4)} = \frac{1}{24} \sum_{\substack{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_5 \\ \boldsymbol{k}_i \neq \boldsymbol{k}_j}}^{\text{uhs}} & A_{\boldsymbol{k}_1} \cdots A_{\boldsymbol{k}_5} A_{\boldsymbol{k}_{12345}} \cos\left(\theta_{\boldsymbol{k}_1} + \theta_{\boldsymbol{k}_2} + \theta_{\boldsymbol{k}_3} + \theta_{\boldsymbol{k}_4} - \theta_{\boldsymbol{k}_5} - \theta_{\boldsymbol{k}_{12345}}\right) \\ & \times p^{(6)}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \boldsymbol{k}_4, -\boldsymbol{k}_5, -\boldsymbol{k}_{12345}),\end{aligned} \tag{5.95}$$

$$\begin{aligned}\mathcal{Q}_3^{(4)} = \frac{1}{36} \sum_{\substack{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_5 \\ \boldsymbol{k}_i \neq \boldsymbol{k}_j}}^{\text{uhs}} & A_{\boldsymbol{k}_1} \cdots A_{\boldsymbol{k}_5} A_{\boldsymbol{k}_{12345}} \cos\left(\theta_{\boldsymbol{k}_1} + \theta_{\boldsymbol{k}_2} + \theta_{\boldsymbol{k}_3} - \theta_{\boldsymbol{k}_4} - \theta_{\boldsymbol{k}_5} - \theta_{\boldsymbol{k}_{12345}}\right) \\ & \times p^{(6)}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, -\boldsymbol{k}_4, -\boldsymbol{k}_5, -\boldsymbol{k}_{12345}),\end{aligned} \tag{5.96}$$

As for the the four-point phase correlator we first integrate out all amplitudes, which now gives rise to a factor $(\sqrt{\pi}/2)^6$ each, where we have made use of Eq. (5.91) again. We then consider a six-point correlator of the form $\langle \epsilon_{\boldsymbol{q}_1} \epsilon_{\boldsymbol{q}_2} \epsilon_{-\boldsymbol{q}_{12}} \epsilon_{\boldsymbol{q}'_1} \epsilon_{\boldsymbol{q}'_2} \epsilon_{-\boldsymbol{q}'_{12}} \rangle$ and assume that $\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}'_1, \boldsymbol{q}'_2 \in$ uhs. In this case only Eq. (5.95) can give a non-vanishing contribution and for that $\boldsymbol{q}_{12}$ and $\boldsymbol{q}'_{12}$ must equal either $\boldsymbol{k}_5$ or $\boldsymbol{k}_{12345}$ and all other $\boldsymbol{q}$- and $\boldsymbol{q}'$-modes must be identified with the remaining $\boldsymbol{k}$-modes. This gives $2 \times 4! = 48$ possibilities and thus we get

$$\langle \epsilon_{\boldsymbol{q}_1} \epsilon_{\boldsymbol{q}_2} \epsilon_{-\boldsymbol{q}_{12}} \epsilon_{\boldsymbol{q}'_1} \epsilon_{\boldsymbol{q}'_2} \epsilon_{-\boldsymbol{q}'_{12}} \rangle = \left(\frac{\sqrt{\pi}}{2}\right)^6 p^{(6)}(\boldsymbol{q}_1, \boldsymbol{q}_2, -\boldsymbol{q}_{12}, \boldsymbol{q}'_1, \boldsymbol{q}'_2, -\boldsymbol{q}'_{12}), \tag{5.97}$$

which is also valid for any configuration of the six modes.

5.A.3 Mixed five-point correlator

Finally, we have to consider the mixed five-point correlator between five phases and two amplitudes, whose leading contributions are of the order $1/V^{3/2}$, i.e. $n = 3$. There are two terms in the Edgeworth expansion, which have five different phase factors:

$$\begin{aligned}\mathcal{Q}_1^{(3)} = \frac{1}{12} \sum_{\substack{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_4 \\ \boldsymbol{k}_i \neq \boldsymbol{k}_j}}^{\text{uhs}} & A_{\boldsymbol{k}_1} \cdots A_{\boldsymbol{k}_4} A_{\boldsymbol{k}_{1234}} \cos\left(\theta_{\boldsymbol{k}_1} + \theta_{\boldsymbol{k}_2} + \theta_{\boldsymbol{k}_3} + \theta_{\boldsymbol{k}_4} - \theta_{\boldsymbol{k}_{1234}}\right) \\ & \times p^{(5)}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \boldsymbol{k}_4, -\boldsymbol{k}_{1234}),\end{aligned} \tag{5.98}$$

$$\mathcal{Q}_2^{(3)} = \frac{1}{6} \sum_{\substack{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_4 \\ \boldsymbol{k}_i \neq \boldsymbol{k}_j}}^{\text{uhs}} A_{\boldsymbol{k}_1} \cdots A_{\boldsymbol{k}_4} A_{\boldsymbol{k}_{1234}} \cos\left(\theta_{\boldsymbol{k}_1} + \theta_{\boldsymbol{k}_2} + \theta_{\boldsymbol{k}_3} - \theta_{\boldsymbol{k}_4} - \theta_{\boldsymbol{k}_{1234}}\right) \times p^{(5)}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, -\boldsymbol{k}_4, -\boldsymbol{k}_{1234}) \,, \tag{5.99}$$

Let us consider the correlator $\langle \delta_{\boldsymbol{k}_1} \delta_{\boldsymbol{k}_2} \epsilon_{\boldsymbol{q}_1} \epsilon_{\boldsymbol{q}_2} \epsilon_{\boldsymbol{q}_3} \rangle$, where we assume that $\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{q}_1, \boldsymbol{q}_2 \in$ uhs and $\boldsymbol{q}_3 \in$ lhs. All terms involving $\mathcal{Q}_2^{(3)}$ will evaluate to zero and from the amplitude integrals we obtain a factor $\mathcal{I}(1)^3 \times \mathcal{I}(2)^2 = (\sqrt{\pi}/2)^3$. We are thus left with the following phase integrals:

$$\begin{aligned} \langle \delta_{\boldsymbol{k}_1} \delta_{\boldsymbol{k}_2} \epsilon_{\boldsymbol{q}_1} \epsilon_{\boldsymbol{q}_2} \epsilon_{\boldsymbol{q}_3} \rangle = \sqrt{P(k_1)\,P(k_2)} \left(\frac{\sqrt{\pi}}{2}\right)^3 \frac{V}{12} \int \prod_{\boldsymbol{p} \in \text{uhs}} \frac{\mathrm{d}\theta_{\boldsymbol{p}}}{2\pi} \\ \times \sum_{\substack{\boldsymbol{u}_1,\ldots,\boldsymbol{u}_4 \in \text{uhs} \\ \boldsymbol{u}_i \neq \boldsymbol{u}_j}} \cos\left(\theta_{\boldsymbol{u}_1} + \cdots + \theta_{\boldsymbol{u}_4} - \theta_{\boldsymbol{u}_{1234}}\right) \\ \times\, p^{(5)}(\boldsymbol{u}_1, \ldots, -\boldsymbol{u}_{1234})\, \mathrm{e}^{i\left(\theta_{\boldsymbol{k}_1} + \boldsymbol{k}_2 + \boldsymbol{q}_2 + \boldsymbol{q}_2 - \theta_{-\boldsymbol{q}_3}\right)} \,, \end{aligned} \tag{5.100}$$

which only give a non-vanishing result if we impose the condition that $\boldsymbol{q}_3 = -\boldsymbol{k}_1 - \boldsymbol{k}_2 - \boldsymbol{q}_1 - \boldsymbol{q}_2$. We then have $4! = 24$ possibilities of matching the various $\boldsymbol{k}$- and $\boldsymbol{q}$-modes with the $\boldsymbol{u}$-vectors and after taking the continuum limit we finally get:

$$\begin{aligned} \langle \delta_{\boldsymbol{k}_1} \delta_{\boldsymbol{k}_2} \epsilon_{\boldsymbol{q}_1} \epsilon_{\boldsymbol{q}_2} \epsilon_{\boldsymbol{q}_3} \rangle = (2\pi)^3 \left(\frac{\sqrt{\pi}}{2}\right)^3 \sqrt{P(k_1)\,P(k_2)}\, p^{(5)}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3) \\ \times\, \delta_D\left(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{q}_1 + \boldsymbol{q}_2 + \boldsymbol{q}_3\right) . \end{aligned} \tag{5.101}$$

As before this result is not restricted to the particular configuration of wavevectors we have chosen above.

Appendix 5.B Galaxy power spectrum at one-loop order

At linear order the power spectrum is only dependent on the single bias parameter b_1, while at the one-loop level non-linear and non-local bias introduce some additional terms that we need to account for. The full galaxy power spectrum at one-loop order is therefore given by McDonald and Roy (2009):

$$\begin{aligned} P_g(k) = b_1^2\, P(k) + 2b_2 b_1\, P_{b2}(k) + 2b_{s^2} b_1\, P_{bs2}(k) + b_2^2\, P_{b22}(k) + 2b_2 b_{s^2}\, P_{b2,bs2}(k) \\ + b_{s^2}^2\, P_{bs22}(k) + 2b_1 b_{3\text{nl}}\, \sigma_3^2(k)\, P_L(k) \,, \end{aligned} \tag{5.102}$$

where $P(k)$ and $P_L(k)$ denote the one-loop and linear dark matter power spectra, respectively. The power spectra that appear in combination with the bias parameters b_2

and b_{s^2} are given by the following integrals,

$$P_{b2}(k) = \int \frac{\mathrm{d}^3 q}{(2\pi)^3} P_L(q)\, P_L(|\boldsymbol{k}-\boldsymbol{q}|)\, F_2(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q})\,, \tag{5.103}$$

$$P_{bs2}(k) = \int \frac{\mathrm{d}^3 q}{(2\pi)^3} P_L(q)\, P_L(|\boldsymbol{k}-\boldsymbol{q}|)\, F_2(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q})\, S_2(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q})\,, \tag{5.104}$$

$$P_{b2,bs2}(k) = -\frac{1}{2}\int \frac{\mathrm{d}^3 q}{(2\pi)^3} P_L(q) \left[\frac{2}{3} P_L(q) - P_L(|\boldsymbol{k}-\boldsymbol{q}|)\, S_2(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q})\right], \tag{5.105}$$

$$P_{bs22}(k) = -\frac{1}{2}\int \frac{\mathrm{d}^3 q}{(2\pi)^3} P_L(q) \left[\frac{4}{9} P_L(q) - P_L(|\boldsymbol{k}-\boldsymbol{q}|)\, S_2(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q})^2\right], \tag{5.106}$$

$$P_{b22}(k) = -\frac{1}{2}\int \frac{\mathrm{d}^3 q}{(2\pi)^3} P_L(q) \left[P_L(q) - P_L(|\boldsymbol{k}-\boldsymbol{q}|)\right], \tag{5.107}$$

$$\sigma_3^2(k) = -\frac{1}{2}\int \frac{\mathrm{d}^3 q}{(2\pi)^3} P_L(q) \left[\frac{5}{6} + \frac{15}{8} S_2(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q})\, S_2(-\boldsymbol{q}, \boldsymbol{k}) - \frac{5}{4} S_2(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q})\right], \tag{5.108}$$

where the kernel functions F_2 and S_2 are defined in Eqs. (5.11) and (5.15). Assuming that galaxy bias is local in Lagrangian space, the non-local terms can be related at first order to the linear bias term as follows (Baldauf et al., 2012; Chan et al., 2012; Saito et al., 2014):

$$b_{s^2} = -\frac{4}{7}(b_1 - 1)\,, \tag{5.109}$$

$$b_{3\mathrm{nl}} = \frac{32}{315}(b_1 - 1)\,. \tag{5.110}$$

Contributions

All of the work presented in this chapter was carried out by myself, including the analytical derivations, as well as the measurements for which I developed an improved version of the code from Chap. 4. I also prepared the figures and wrote the complete manuscript, which got published in Eggemeier and Smith (2017). My supervisor (co-author on this paper) gave guidance on the overall project, provided the data of the N-body simulations along with appropriate read-in routines, and made minor edits to the draft.

Chapter 6

Towards optimal cosmological parameter recovery from compressed bispectrum statistics

One of the underlying motivations for studying the LCF in Chaps. 4 and 5 was its ability to probe information highly complementary to the power spectrum. This stands in contrast with the bispectrum, which measures a combination of Fourier amplitudes and phases, and is therefore expected to be more correlated with traditional two-point statistics. This leads us to the question whether the LCF is a more efficient probe of the three-point statistics regime, in the sense that it extracts the same information than the bispectrum, but from a smaller number of configurations. If true, this would improve on the covariance matrix problematic as outlined in Sec. 3.3.2. The present chapter aims to investigate this possibility by comparing the parameter sensitivity of the LCF and bispectrum for the same set of parameters used in Chap. 5. In addition, we include and compare with two further proxies of the bispectrum: the integrated bispectrum and a modal decomposition. We find that the latter provides the best performance and reduces the size of the bispectrum covariance matrix by at least an order of magnitude without losing any information. The LCF, on the other hand, is unlikely to yield a lossless compression.

6.1 Introduction

Constraints on cosmological parameters have improved significantly over the last two decades, driven by high-precision data from the cosmic microwave background ('CMB') temperature and polarisation anisotropies (Bennett et al., 2013; Planck Collaboration et al., 2016b). But the capacity of CMB observations to sustain this rate of progress is now nearly exhausted. Measurements of the temperature anisotropy have become limited

by cosmic variance down to very small scales, and therefore future large-scale measurements will furnish little new information. Meanwhile, on small scales, cosmological information begins to be erased by astrophysical processes. Modest improvements may still come from better polarisation data, perhaps shrinking current uncertainties by a factor of a few, but eventually these measurements will also approach the limit of cosmic variance. Further progress will be possible only with new sources of information. In the decade 2020–2030 we expect such a source to be provided by surveys of cosmological large-scale structure—but only if the information these surveys contain can be extracted and understood (Silk, 2017).

6.1.1 The bispectrum: challenges

The statistical information contained in a galaxy survey is carried by the probability density of the fluctuations, which can be measured using a hierarchy of suitably-chosen correlation functions, perhaps involving a transformation of the density field (Carron and Szapudi, 2013; Neyrinck et al., 2009; Seo et al., 2011). Tools to extract information from the two-point function were developed early and are now mature. The development of tools to extract information from higher-order functions has proceeded more slowly (Fry, 1984; Goroff et al., 1986; Scoccimarro, 2000; Sefusatti et al., 2006), but because structure formation is non-linear it is likely that these carry an important fraction of the information content. To make good use of our investment in costly observational programmes it will be necessary to find a means of using information from at least the three-point function. So far, there are few bispectrum (or 3-point function) measurements from modern surveys compared to power spectrum analyses (Gil-Marín et al., 2015, 2017; Marin et al., 2013; Slepian et al., 2017a,b).

What are the challenges? A first difficulty arises from combinatorics. We write the matter over-density at time t as $\delta(\boldsymbol{x},t) = \delta\rho(\boldsymbol{x},t)/\bar{\rho}(t)$, where $\delta\rho(\boldsymbol{x},t) = \rho(\boldsymbol{x},t) - \bar{\rho}(t)$ is the density perturbation and $\bar{\rho}(t)$ is the uniform background. Allowing angle brackets $\langle\cdots\rangle$ to denote an ensemble average, statistical homogeneity makes the two- and three-point functions $\langle\delta(\boldsymbol{x})\delta(\boldsymbol{x}+\boldsymbol{r})\rangle$ and $\langle\delta(\boldsymbol{x})\delta(\boldsymbol{x}+\boldsymbol{r}_1)\delta(\boldsymbol{x}+\boldsymbol{r}_2)\rangle$ independent of the origin $\boldsymbol{x}$. After translation to Fourier space this enforces conservation of momentum for the wave numbers that participate in the expectation value,

$$\langle\delta(\boldsymbol{k}_1)\delta(\boldsymbol{k}_2)\rangle = (2\pi)^3\delta_{\mathrm{D}}(\boldsymbol{k}_1+\boldsymbol{k}_2)P(k), \tag{6.1a}$$

$$\langle\delta(\boldsymbol{k}_1)\delta(\boldsymbol{k}_2)\delta(\boldsymbol{k}_3)\rangle = (2\pi)^3\delta_{\mathrm{D}}(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)B(k_1,k_2,k_3), \tag{6.1b}$$

where $k = |\boldsymbol{k}_1| = |\boldsymbol{k}_2|$ is the common magnitude of the wave numbers appearing in the two-point function. In Eqs. (6.1a)–(6.1b) and the remainder of this chapter we suppress the time t labelling the hypersurface of evaluation. Isotropy makes the power

spectrum P a function only of k, while the bispectrum B is a function of the three wave numbers k_1, k_2, k_3 subject to the closure condition $\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3 = 0$. Therefore a fixed volume of space yields many more distinct configurations of the bispectrum than of the spectrum. If we choose to measure all of them then we must provide an estimate for their covariance, and beyond the Gaussian approximation this typically requires N-body simulations. Since we require at least as many simulations as the number of independent covariances, the number of simulations to be performed grows at least linearly in the number of configurations. This makes it very expensive to use more than a fraction of the available bispectrum measurements.

Second, we must estimate typical values for $B(k_1, k_2, k_3)$ in a particular cosmological model. While such estimates are already necessary for the power spectrum $P(k)$, accurate estimates for the bispectrum are substantially more challenging. There are two key reasons. No matter what methods we use, the algebraic complexity associated with high-order correlation functions is usually worse than at lower order. Also, many of our standard tools have a reduced range of validity as we move up the correlation hierarchy. We must therefore work harder to obtain trustworthy predictions from our models, and in some cases we can do so only by giving up analytic methods altogether.

These problems have hampered the development of a toolkit that would make use of bispectrum measurements routine. Nevertheless, they are difficulties of practice and not obstructions of principle—if necessary, we could determine both covariances and typical values of P or B from N-body simulations, at least over a certain range of scales. But such determinations would require a very large number of realisations. The sheer computational resource entailed by this strategy makes it unattractive on timescales of interest for surveys such Euclid, DESI, or LSST.

6.1.2 Alternative strategies

To build a practical methodology we must cut the size of the covariance matrices and avoid simulations where possible. Simulations are not needed when analytic methods suffice to predict P or B, or when a Gaussian approximation to the covariance is acceptable. Meanwhile, an obvious way to reduce the number of configurations is simply not to measure them all. Depending how aggressively we choose to cut, this may mean accepting a significant loss of information. A more nuanced option is to aggregate groups of configurations into weighted averages, effectively *compressing* the data carried by the bispectrum rather than discarding it. Such averages could be computed directly. But there are also observables whose statistics can naturally be expressed as weighted averages of this kind. Measuring these will often be simpler than measuring amplitudes of the Fourier bispectrum—simultaneously reducing the effort required to estimate and

invert their covariance matrices. We describe these observables as 'proxies' or 'proxy statistics' for the full Fourier bispectrum.

Each proxy represents a compromise between (*a*) information loss due to compression, (*b*) the type of Fourier configurations over which it aggregates, and therefore the physics to which it is sensitive, and (*c*) its accessibility to analytical modelling, either for covariances or to estimate typical measurements. In this chapter we select three proxies that have already been described in the literature and characterise their performance in each of these categories. Our aim is not to find an optimal proxy for any particular measurement, but rather to demonstrate that their use represents a feasible strategy for upcoming surveys without unacceptable degradation in information recovery.

6.1.3 Summary

Our principal results are forecasts for the parameter error bars achievable from combinations of the galaxy power spectrum and bispectrum, or its proxies. The parameter set we study comprises the background quantities of a ΛCDM model with evolving dark energy, supplemented by two parameters describing the bias model (McDonald and Roy, 2009). We study how these forecasts change when they are estimated using the complete non-Gaussian covariance matrix or its Gaussian approximation. We characterise their dependence on the method used to predict typical values for $P(k)$ and $B(k_1, k_2, k_3)$ by sampling the results using tree-level and one-loop standard perturbation theory ('SPT'), and an implementation of the halo model. We compare these estimates with values measured directly from simulations. These results can be used to determine, for each observable, the degree of modelling sophistication that is required to obtain accurate forecasts.

Our analysis does not include the effect of survey geometry or incompleteness, or redshift-space effects, and should be regarded as a determination of the performance of each proxy under idealised conditions. We include a simple analysis that indicates how our results would change in the presence of shot noise.

Fisher forecasts including Fourier bispectrum measurements have previously been reported by Sefusatti et al. (2006), assuming 1,015 bispectrum configurations and measuring covariances from a suite of 6,000 mock catalogues generated by the **PTHalos** algorithm (Scoccimarro and Sheth, 2002) and second-order Lagrangian perturbation theory ('2LPT'). Their results suggested that the bispectrum contains significant cosmological information. For comparison, in our analysis we use 95 bispectrum configurations in order to keep the size of the covariance matrix within plausible bounds, and measure it directly from a suite of full N-body simulations.

More recently, Chan and Blot (2017) estimated the extra constraining power of Fourier bispectrum measurements by computing their contribution to the signal-to-

noise, but did not make forecasts for error bars on cosmological parameters. They concluded that the bispectrum contributes a modest increase in the signal-to-noise above the power spectrum, with the additional information perhaps being principally useful to break degeneracies. One of our aims is to clarify the relationship between this conclusion and the more nuanced outcomes found by Sefusatti et al. (2006). We find that estimates based on signal-to-noise alone generally give only a rough indication compared to the full Fisher calculation because they do not account for variations in the sensitivity to background cosmology between observables.

6.1.4 Organisation

Our presentation is organised as follows. In Sec. 6.2 we introduce the three bispectrum proxies to be studied in the remainder of the chapter. These are: (*a*) the *integrated bispectrum* (Chiang et al., 2014), which measures variation of the power spectrum in subsampled regions ; (*b*) the *line correlation function*, which samples three-point statistics of the phase of the density fluctuation (Obreschkow et al., 2013; Wolstenhulme et al., 2015), and (*c*) the *modal bispectrum*, which can be regarded as an alternative to the Fourier bispectrum obtained by exchanging the Fourier modes $e^{i\boldsymbol{k}\cdot\boldsymbol{x}}$ for an alternative basis (Fergusson et al., 2012; Regan et al., 2012). Each of these measures can be expressed as a weighted average over particular configurations of the Fourier bispectrum.

In Secs. 6.3.2–6.3.4 we explain how each proxy can be predicted using the halo model or a flavour of SPT. In Sec. 6.3.5 we explain our prescription to obtain the biased galaxy density field from the underlying matter density field, which is the quantity predicted by these analytic models. In Sec. 6.4 we describe our procedure to recover estimates for each proxy statistic from N-body simulations, and in Sec. 6.5 we compare these estimates (and estimates for their derivatives with respect to the cosmological parameters) with theoretical predictions. Readers familiar with the measures of 3-point correlations described in Sec. 6.2 and the modelling technologies of Sec. 6.3 may choose to begin reading at this point. In Sec. 6.6 we present signal-to-noise estimates for the information content of each proxy. Our Fisher forecasts appear in Sec. 6.7. In Sec. 6.8 we collect a number of topics for discussion, including the compression efficiency of each proxy statistic and the impact of shot noise on our forecasts. We conclude in Sec 6.9.

6.2 The Fourier bispectrum and its proxies

In this section we introduce the proxy statistics to which we compare the Fourier bispectrum. This has already been defined—together with the power spectrum—in Eqs. (6.1a)–(6.1b). We describe the integrated bispectrum in Sec. 6.2.1, the line correlation function in Sec. 6.2.2 and the modal decomposition of the bispectrum in Sec. 6.2.3. Each of these

represents a possible compression of the Fourier bispectrum, in the sense described in Sec. 6.1.

6.2.1 Integrated bispectrum

The integrated bispectrum (or 'position-dependent power spectrum') was developed by Chiang et al. (2014) as a tool to search for primordial non-Gaussianity in large-scale structure. It has several convenient features: it is easily estimated using standard power-spectrum codes and it has a clear physical interpretation. As we shall see in Sec. 6.3.2, it represents a weighted average of the Fourier bispectrum dominated by 'squeezed' configurations — that is, wave numbers $(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$ where one k_i is much smaller than the other two. If we assume $k_3 \ll k_1, k_2$ then the bispectrum $\langle\delta(\boldsymbol{k}_1)\delta(\boldsymbol{k}_2)\delta(\boldsymbol{k}_3)\rangle$ expresses correlations between a single long-wavelength mode $\delta(\boldsymbol{k}_3)$ and the two-point function $\langle\delta(\boldsymbol{k}_1)\delta(\boldsymbol{k}_2)\rangle$. This makes it sensitive to 'local-type' non-Gaussianity produced by inflationary models with more than one active field. However, because gravitational collapse correlates modes with comparable wave numbers, the bispectrum produced during mass assembly is typically concentrated away from squeezed configurations. For this reason it is not clear how sensitive the integrated bispectrum might be to the cosmological parameters that influence this assembly process.

To define the integrated bispectrum divide the total survey volume into N_s cubic subvolumes, each of volume $V_s \equiv L_s^3$ and centred at positions $\boldsymbol{r}_L$. Compute the power spectrum and average over-density for each subvolume, which we denote $P(\boldsymbol{k}, \boldsymbol{r}_L)$ and $\bar{\delta}(\boldsymbol{r}_L)$, respectively. (The power spectrum $P(\boldsymbol{k}, \boldsymbol{r}_L)$ may depend on the orientation of $\boldsymbol{k}$ if the subvolumes are not isotropic.) Finally, the integrated bispectrum is defined to be the expectation of $P(\boldsymbol{k}, \boldsymbol{r}_L)\bar{\delta}(\boldsymbol{r}_L)$, averaged over the orientation of $\boldsymbol{k}$,

$$iB(k) \equiv \int \frac{\mathrm{d}^2\hat{k}}{4\pi} \langle P(\boldsymbol{k}, \boldsymbol{r}_L)\bar{\delta}(\boldsymbol{r}_L)\rangle_{N_s}\,. \tag{6.2}$$

The notation $\langle\cdots\rangle_{N_s}$ indicates that the expectation is to be taken over all subvolumes.

To compute this expectation we Taylor expand $P(\boldsymbol{k}, \boldsymbol{r}_L)$ in powers of $\bar{\delta}(\boldsymbol{r}_L)$ (Chiang et al., 2014). The leading contribution is

$$\langle P(\boldsymbol{k}, \boldsymbol{r}_L)\bar{\delta}(\boldsymbol{r}_L)\rangle_{N_s} = \left\langle \left[P(\boldsymbol{k})\big|_{\bar{\delta}=0} + \frac{\mathrm{d}P(\boldsymbol{k})}{\mathrm{d}\bar{\delta}}\Big|_{\bar{\delta}=0} \bar{\delta}(\boldsymbol{r}_L) + \cdots \right] \bar{\delta}(\boldsymbol{r}_L) \right\rangle_{N_s} \approx \frac{\mathrm{d}\ln P(\boldsymbol{k})}{\mathrm{d}\bar{\delta}}\Big|_{\bar{\delta}=0} P(\boldsymbol{k})\sigma_L^2\,, \tag{6.3}$$

where $\sigma_L^2 \equiv \langle\bar{\delta}^2(\boldsymbol{r}_L)\rangle_{N_s}$ is the variance in mean over-density over the subvolumes. Therefore, at lowest order, the integrated bispectrum describes variation of the power spectrum in response to changes in the large-scale over-density.[1] We conclude that measurements of iB contain both the power spectrum and its variance. Since these can be

[1] In field theory this is the 'operator product expansion'.

measured directly, any new information contained in the integrated bispectrum must reside in its normalised component (Chiang et al., 2014),

$$ib(k) \equiv \frac{iB(k)}{P(k)\sigma_L^2} \approx \left.\frac{\mathrm{d}\ln P(k)}{\mathrm{d}\bar{\delta}}\right|_{\bar{\delta}=0}, \tag{6.4}$$

where the second approximate equality applies when only the lowest-order contribution from the Taylor expansion need be retained. This is the linear response approximation. The quantity $\mathrm{d}\ln P(k)/\mathrm{d}\bar{\delta}$ is the *linear response function* and provides a good approximation to ib for large k.

6.2.2 Line Correlation Function

Eq. (6.1a) shows that the power spectrum is sensitive only to information carried by the amplitude of each Fourier mode. In contrast, higher-order statistics generally encode information carried by both amplitudes and phases. Phase correlations are an exclusive signature of non-Gaussian density fields. For instance, they may arise through processes in the primordial Universe or from mode coupling in the non-linear regime of gravitational collapse. Therefore, unlike the amplitudes, phases directly probe cosmological information that is absent from the two-point function.

With this motivation, (Obreschkow et al., 2013) proposed the line correlation function (often abbreviated as 'LCF'). It measures a subset of three-point phase correlations of the density field—specifically, correlations between collinear points, each separated by a distance r. (Obreschkow et al., 2013) demonstrated that the LCF is a robust tracer of filamentary structures, and showed that it could be used as a phenomenological tool to distinguish between cold and warm dark matter scenarios. Subsequent work established its connection to conventional higher-order statistics (Eggemeier et al., 2015; Eggemeier and Smith, 2017; Wolstenhulme et al., 2015).

The line correlation function can be understood as follows: for a given density field $\delta(\boldsymbol{x})$ in some volume V, its real-space phase field $\epsilon_r(\boldsymbol{x})$ smoothed on a scale r satisfies

$$\epsilon_r(\boldsymbol{x}) = \int \frac{\mathrm{d}^3k}{(2\pi)^3}\epsilon(\boldsymbol{k})\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}W(k|r) \equiv \int \frac{\mathrm{d}^3k}{(2\pi)^3}\frac{\delta(\boldsymbol{k})}{|\delta(\boldsymbol{k})|}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}W(k|r), \tag{6.5}$$

where $W(k|r)$ is the Fourier transform of the smoothing window function. We take this to be a spherical top-hat in k-space, $W(k|r) \equiv \Theta(1 - k\,r/2\pi)$, where $\Theta(x)$ denotes the Heaviside step function. The phase at $\boldsymbol{k} = 0$ is defined so that $\epsilon(0) \equiv 0$. Following Obreschkow et al. (2013), the LCF is defined by

$$\ell(r) \equiv \frac{V^3}{(2\pi)^9}\left(\frac{r^3}{V}\right)^{3/2}\int \frac{\mathrm{d}^2\hat{r}}{4\pi}\left\langle\epsilon_r(\boldsymbol{x})\epsilon_r(\boldsymbol{x}+\boldsymbol{r})\epsilon_r(\boldsymbol{x}-\boldsymbol{r})\right\rangle, \tag{6.6}$$

where the factor $V^3/(2\pi)^9$ represents a volume regularisation. After taking Fourier transforms we require the three-point function of the $\epsilon_r(\boldsymbol{k})$ in order to evaluate this integral. Wolstenhulme et al. (2015) and Eggemeier and Smith (2017) demonstrated that, at lowest order in the expansion of the probability density function for Fourier phases, this three-point function is directly related to the Fourier bispectrum. Therefore the LCF must contain some fraction of the information in B, but because $\ell(r)$ is an average over specific collinear configurations it represents a compression. Specifically, the number of LCF bins will vary linearly with changes in the effective cut-off on Fourier modes.

6.2.3 Modal bispectrum

Our final proxy is a 'modal' expansion of the three-point function. This is very similar to the Fourier bispectrum, except that we exchange the Fourier basis $e^{i\boldsymbol{x}\cdot\boldsymbol{k}}$ for a set of alternative modes that are better adapted to the structure of B. The exchange is helpful if we can represent the bispectrum to the same accuracy using fewer modes than required by the Fourier representation. This approach was originally developed by Fergusson and Shellard (2009) and Regan et al. (2010) to analyse microwave background data, and subsequently applied to large-scale structure by Fergusson, Regan, and Shellard (2012) and Regan et al. (2012).

In the alternative basis we represent the Fourier bispectrum in the form

$$B(k_1,k_2,k_3) \approx B_{\rm modal}(k_1,k_2,k_3) \equiv \frac{1}{w(k_1,k_2,k_3)} \sum_{n=0}^{n_{\rm max}-1} \beta_n^Q Q_n(k_1,k_2,k_3), \tag{6.7}$$

where the Q_n are basis functions that span the space of configurations compatible with a triangle condition on (k_1,k_2,k_3), but can otherwise be chosen freely provided they are linearly independent. The β_n^Q are numbers that we describe as 'modal coefficients'. They can be regarded as averages of the Fourier bispectrum over a set of configurations picked out by the corresponding Q_n. The function $w(k_1,k_2,k_3)$ is an arbitrary weight that will be chosen in Sec. 6.3.4.

If the Q_n form a complete basis we expect B and $B_{\rm modal}$ to become equivalent in the limit $n_{\rm max}\to\infty$. In this limit the modal expansion is merely a reorganisation of the Fourier representation. But if we select the lowest Q_n to average over the most relevant Fourier configurations then it may be possible to represent a typical B using only a small number of modes.[2] Taking $n_{\rm max}$ to be of order this number, the outcome yields useful compression whenever $n_{\rm max} \ll N_{\rm triangles}$, where $N_{\rm triangles}$ is the number of

[2]Here, 'most relevant' is defined by the features of the bispectrum for which we wish to search. For example, inspection of the formulae appearing in Sections 6.3.2–6.3.3 below shows that both the integrated bispectrum and line correlation function can be regarded as instances of (6.7), with Q_n adjusted to prioritise specific groups of Fourier configurations. For these cases, however, the resulting Q-basis is not complete. In this chapter we distinguish the modal decomposition, for which the Q-basis is intended to be complete, from proxies such as ib and ℓ which are intended to be projections.

Fourier configurations contained in the volume under discussion. At least for reasonably smooth bispectra, Schmittfull, Regan, and Shellard (2013) found that this could be done with no more than modest loss of signal.

Orthonormal basis

Given a choice of Q_n we may redefine the basis by taking arbitrary linear combinations. For example, we will use this freedom in Sec. 6.3.4 to obtain a basis for which the β-coefficients are uncorrelated. The covariance matrix in this redefined basis is especially simple.

Such a redefinition can be performed using an invertible matrix λ_{mn}. We define $R_n \equiv \sum_m \lambda^{-1}_{nm} Q_m$. The β-coefficients in the R-basis now satisfy $\beta^R_n \equiv \sum_m \lambda_{mn} \beta^Q_m$. Since the Q- and R-bases are reorganisations of each other, the modal bispectrum defined using either basis is equivalent,

$$B(k_1,k_2,k_3) \approx \frac{1}{w(k_1,k_2,k_3)} \sum_{n=0}^{n_{\max}-1} \beta^Q_n Q_n(k_1,k_2,k_3) = \frac{1}{w(k_1,k_2,k_3)} \sum_{n=0}^{n_{\max}-1} \beta^R_n R_n(k_1,k_2,k_3). \tag{6.8}$$

6.3 Predicting typical values and covariances for the proxies

In this section we explain how to obtain predictions for the typical values and covariances of $ib(k)$, $\ell(r)$ and β^R_m in a given cosmological model. This can be done with different degrees of sophistication, corresponding — for example — to truncations at different levels in the loop expansion of standard perturbation theory (Bernardeau et al., 2002), or by using fitting functions calibrated to match the output of N-body simulations (Mead et al., 2015). Since each proxy aggregates a different group of Fourier configurations, and these configurations vary in their response to features of the background cosmology, the sophistication needed to adequately capture the behaviour of the proxies may vary.

This is both a challenge and an opportunity. Proxies that require delicate modelling to obtain accurate predictions are harder to use, and may be expensive to deploy in a parameter-estimation Monte Carlo. In favourable cases, however, the payoff will be sensitive discrimination between nearby cosmological models. On the other hand, proxies that can be modelled robustly using simple methods are easy to use and cheap to deploy, but may offer correspondingly coarse discrimination. We study these trade-offs by contrasting predictions made using tree-level and one-loop SPT, and the halo model. For the halo-model power spectrum we choose the HMcode implementation that was originally described in Mead et al. (2015). For the halo-model bispectrum we

use the standard formulae given by Cooray and Sheth (2002) with a Sheth–Tormen mass function (Sheth and Tormen, 1999) and Navarro–Frenk–White halo profile (Navarro et al., 1996), unless the text indicates differently. In Sec. 6.5 we study the performance of each method compared to numerical estimates extracted directly from N-body simulations, which enables us to characterise the minimum adequate sophistication for each proxy. For simplicity our analysis is framed in terms of the underlying dark matter density field, although in Sec. 6.3.5 we explain how this can be extended to predict galaxy clustering.

6.3.1 Covariance

To compute a likelihood for a given proxy, either for the purposes of parameter estimation or to make forecasts, we require an estimate for the covariance between different configurations. Therefore the minimum sophistication needed to adequately predict this covariance matrix will play an additional role in determining the relative expense of each proxy. In practice the covariance matrix is typically estimated by taking measurements from a large suite of N-body simulations or 2LPT catalogues, or, if this is cannot be done, by falling back to a Gaussian approximation. N-body simulations give accurate results, but are expensive enough that assembling sufficient independent realisations to determine the inverse covariance is often not feasible. In comparison, catalogues based on 2LPT are significantly cheaper but become inaccurate in the non-linear regime, while the Gaussian prediction breaks down even earlier and may miss cross-correlations that significantly affect the outcome.

The relative importance of these cross-correlations varies between proxies. In Sections 6.6–6.7 we estimate their significance by comparing results from N-body and Gaussian covariances. We describe our procedure to estimate covariance matrices from the simulations in Sec. 6.5, but collect formulae for the Gaussian approximation here.

For comparison, the Gaussian covariance for the power spectrum and Fourier bispectrum, measured on a grid of spacing Δk with fundamental frequency $k_\mathrm{f} = 2\pi / V^{1/3}$, can be written

$$\mathrm{Cov}_\mathrm{G}[P(k_i), P(k_j)] \approx \mathbf{1}_{ij} \frac{2k_\mathrm{f}^3}{4\pi k_i^2 \Delta k} P^2(k_i), \tag{6.9}$$

where $\mathbf{1}_{ij}$ is the Kronecker symbol, and

$$\mathrm{Cov}_\mathrm{G}[B(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3), B(\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3)] \approx \mathbf{1}_{\boldsymbol{k},\boldsymbol{q}} \frac{\mathsf{N}\pi k_\mathrm{f}^3}{k_1 k_2 k_3 (\Delta k)^3} P(k_1) P(k_2) P(k_3). \tag{6.10}$$

The Kronecker symbol $\mathbf{1}_{\boldsymbol{k},\boldsymbol{q}}$ should be interpreted to equal unity if the triangles defined by $\{\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3\}$ and $\{\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3\}$ are equal, and zero otherwise. The degeneracy factor N equals unity for a scalene triangle, two for an isosceles triangle and six for an equilateral triangle.

6.3.2 Integrated bispectrum

To evaluate the expression (6.4) we first establish its relation to the underlying 3-point function. The over-density within the subvolume labelled by $\boldsymbol{r}_L$ can be written

$$\delta(\boldsymbol{k},\boldsymbol{r}_L) = \int \frac{\mathrm{d}^3 q}{(2\pi)^3}\,\delta(\boldsymbol{k}-\boldsymbol{q})\,W_L(\boldsymbol{q})\mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}_L}, \tag{6.11}$$

where $W_L(\boldsymbol{q}) = V_s\prod_{i=1}^3 \mathrm{sinc}(q_i L_s/2)$ is the Fourier transform of the cubic window function with side length L_s, and $\mathrm{sinc}\,x \equiv (\sin x)/x$. The power spectrum in this subvolume is $P(\boldsymbol{k},\boldsymbol{r}_L) \equiv \langle|\delta(\boldsymbol{k},\boldsymbol{r}_L)|^2\rangle/V_s$ and the mean over-density is $\bar{\delta}(\boldsymbol{r}_L) \equiv \delta(0,\boldsymbol{r}_L)/V_s$. Combining these with Eq. (6.2) yields (Chiang et al., 2014)

$$\begin{aligned} iB^{\mathrm{theory}}(k) = \frac{1}{V_s^2}\int\frac{\mathrm{d}^2\hat{k}}{4\pi}\int\frac{\mathrm{d}^3 q_1}{(2\pi)^3}\int\frac{\mathrm{d}^3 q_2}{(2\pi)^3}\,B^{\mathrm{theory}}(\boldsymbol{k}-\boldsymbol{q}_1,-\boldsymbol{k}+\boldsymbol{q}_1+\boldsymbol{q}_2,-\boldsymbol{q}_2) \\ \times W_L(\boldsymbol{q}_1)W_L(-\boldsymbol{q}_1-\boldsymbol{q}_2)W_L(\boldsymbol{q}_2)\,. \end{aligned} \tag{6.12}$$

Because $\mathrm{sinc}\,x$ is strongly peaked for $|x| \lesssim \pi$ the window functions W_L effectively constrain the q_i integrals to $q_i \lesssim 1/L_s$. Since $k \gtrsim 1/L_s$ within each subvolume, the integral receives significant contributions only from squeezed configurations of the Fourier bispectrum that are of order the subvolume size or larger, because in the limit $q_1, q_2 \ll k$ we have $B^{\mathrm{theory}}(\boldsymbol{k}-\boldsymbol{q}_1,-\boldsymbol{k}+\boldsymbol{q}_1+\boldsymbol{q}_2,-\boldsymbol{q}_2) \approx B^{\mathrm{theory}}(\boldsymbol{k},-\boldsymbol{k},-\boldsymbol{q}_2)$.

Chiang et al. (2014) computed the linear response function using (6.12) and tree-level SPT, and verified that it reproduces Eq. (6.4) to within 2% for $k \gtrsim 0.2\,h\,\mathrm{Mpc}^{-1}$. For our purposes we require accurate estimates at smaller k, and therefore we perform a numerical integration using (6.12) directly. The integral is 8-dimensional and its evaluation is challenging; we implement it using the **Vegas** algorithm provided by the **CUBA** package (Hahn, 2016). To make the integration time feasible we densely sample B^{theory} on a 3-dimensional cubic mesh in coordinates (k_1, k_2, μ_{12}), where $\mu_{12} \equiv (k_1^2+k_2^2-k_3^2)/(2k_1k_2)$ is the cosine of the angle between $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ and can be used in place of the third wave number k_3. We construct a 3-dimensional cubic spline that interpolates between lattice points and use this spline to evaluate the integrand. To validate this procedure we have verified that our numerical results match the analytic prediction from the linear response function at large k.

Although we have not written subvolume labels explicitly, σ_L^2 and all power spectra in (6.4) refer to subsampled quantities, and therefore should be computed by appropriate convolution with the subvolume window function $W_L(\boldsymbol{q})$.

Halo model

Evaluation of Eq. (6.12) yields good results for tree-level and one-loop SPT, but does not perform well when B^{theory} is given by the standard halo model expression that is

described in Cooray and Sheth (2002). In this case we do not recover equivalence between (6.12) and the linear response function (Eq. 6.4), which we obtain from computing the derivative $\mathrm{d}\ln P/\mathrm{d}\bar{\delta}$ using the simulation-calibrated formula proposed by Chiang et al. (2014),

$$\frac{\mathrm{d}\ln P^{\mathrm{halo}}(k)}{\mathrm{d}\bar{\delta}} = \frac{13}{21}\frac{\mathrm{d}\ln P^{\mathrm{halo}}(k)}{\mathrm{d}\ln\sigma_8} + 2 - \frac{1}{3}\frac{\mathrm{d}\ln k^3 P^{\mathrm{halo}}(k)}{\mathrm{d}\ln k}, \tag{6.13}$$

and numerical differentiation of the HMcode power spectrum. The difference between (6.12) and the linear response function is 3% to 7% for $k \gtrsim 0.2\,h\,\mathrm{Mpc}^{-1}$. We interpret this disagreement as an indication that the standard halo model makes inconsistent predictions for the modulation of the power spectrum with $\bar{\delta}$, or the squeezed limit of the bispectrum, or both. Moreover, comparison of the halo-model ib computed using (6.12) to our N-body simulations shows poor agreement, whereas Eq. (6.13) agrees with our simulations to within 2% for $k \gtrsim 0.2\,h\,\mathrm{Mpc}^{-1}$, suggesting that estimates based on (6.12) will be inaccurate. Therefore, for the halo model only, we estimate ib by assuming the linear response approximation.

Covariance

In the absence of shot noise, the Gaussian covariance for estimates of ib constructed from data can be written

$$\mathrm{Cov}_{\mathrm{G}}\left[ib(k_i), ib(k_j)\right] = \frac{V_s}{V N_{ks}}\frac{1}{\sigma_L^2}\mathbf{1}_{ij}. \tag{6.14}$$

In this expression V_s is the volume of a subsampled region and V denotes the total survey volume. The quantity $N_{ks} = 2\pi k^2 \Delta k V_s$ is the number of Fourier modes in a subvolume k-bin.

6.3.3 Line correlation function

Wolstenhulme et al. (2015) used tree-level SPT to predict the line correlation function. Their result was generalised to an arbitrary bispectrum by Eggemeier and Smith (2017), who gave the formula

$$\ell^{\mathrm{theory}}(r) \simeq \left(\frac{r}{4\pi}\right)^{9/2} \iint\limits_{\substack{|\boldsymbol{k}_1|,\,|\boldsymbol{k}_2|,\\ |\boldsymbol{k}_1+\boldsymbol{k}_2|\,\leq\,2\pi/r}} \mathrm{d}^3k_1\,\mathrm{d}^3k_2\, B_\epsilon^{\mathrm{theory}}(k_1,k_2,k_3)\, j_0\big(|\boldsymbol{k}_1-\boldsymbol{k}_2|\,r\big), \tag{6.15}$$

where $j_0(x) = \sin(x)/x$ is the spherical Bessel function of order zero and the integrals over $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ are cut off at the scale $k_i = 2\pi/r$. The quantity B_ϵ is defined by

$$B_\epsilon(k_1,k_2,k_3) \equiv \frac{B(k_1,k_2,k_3)}{\sqrt{P(k_1)P(k_2)P(k_3)}} \tag{6.16}$$

and gives the dominant contribution to the bispectrum of the phase field $\epsilon(\boldsymbol{k}) = \delta(\boldsymbol{k})/|\delta(\boldsymbol{k})|$ in the limit of large volume V,

$$\langle \epsilon(\boldsymbol{k}_1)\,\epsilon(\boldsymbol{k}_2)\,\epsilon(\boldsymbol{k}_3)\rangle = \frac{(2\pi)^3}{V^{3/2}}\left(\frac{\sqrt{\pi}}{2}\right)^3 B_\epsilon(k_1,k_2,k_3)\,\delta_D(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)\,, \tag{6.17}$$

while for smaller volumes there are corrections scaling as powers of $V^{-1/2}$ compared to the dominant term (Eggemeier and Smith, 2017). In the following we will therefore refer to B_ϵ as the *phase bispectrum*.

Evaluation

To evaluate (6.15) we must perform a 6-dimensional integral. We use a strategy similar to that described in Sec. 6.3.2, by sampling the bispectrum over a cubic lattice and interpolating between lattice sites. The integration is again performed using Vegas.

In the special case of tree-level SPT, Wolstenhulme et al. (2015) showed that (6.15) could be reduced to a 3-dimensional integral,

$$\begin{aligned}\ell^{\rm tree}(r) = 16\pi^2\left(\frac{r}{4\pi}\right)^{9/2}\int_0^{\frac{2\pi}{r}}\mathrm{d}k_1\,k_1^2\int_0^{\frac{2\pi}{r}}\mathrm{d}k_2\,k_2^2\int_{-1}^{\mu_{\rm cut}}\mathrm{d}\mu\,F_2^{(s)}(k_1,k_2,\mu)\sqrt{\frac{P^{\rm tree}(k_1)P^{\rm tree}(k_2)}{P^{\rm tree}(|\boldsymbol{k}_1+\boldsymbol{k}_2|)}}\\ \times\left[j_0\big(|\boldsymbol{k}_2-\boldsymbol{k}_1|\,r\big)+2\,j_0\big(|\boldsymbol{k}_1+2\boldsymbol{k}_2|\,r\big)\right],\end{aligned} \tag{6.18}$$

where $P^{\rm tree}$ is the tree-level power spectrum, and

$$F_2^{(s)}(k_1,k_2,\mu) = \frac{5}{7}+\frac{\mu}{2}\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right)+\frac{2}{7}\mu^2 \tag{6.19}$$

is the symmetrised second order SPT (Bernardeau et al., 2002) kernel with μ being the angle between $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$. The upper limit of the μ-integral is chosen to guarantee $|\boldsymbol{k}_1+\boldsymbol{k}_2| \le 2\pi/r$, which requires

$$\mu_{\rm cut} = \min\left\{1, \max\left\{-1, \frac{(2\pi/r)^2-k_1^2-k_2^2}{2k_1k_2}\right\}\right\}. \tag{6.20}$$

Eq. (6.18) is useful because it provides a means to test the accuracy of our 6-dimensional Vegas integrations, and the 3-dimensional interpolations they entail. We have compared estimates for the tree-level line correlation function using both (6.15) and (6.18) and find that the agreement is better than 0.1 % at the smallest r bin considered and worsens to only ~ 2 % at the largest one.

Covariance

To determine the Gaussian covariance we require the two-point function of the phase field,

$$\langle \epsilon(\boldsymbol{k}_1)\epsilon(\boldsymbol{k}_2)\rangle = \frac{(2\pi)^3}{V}\delta_{\mathrm{D}}(\boldsymbol{k}_1+\boldsymbol{k}_2). \tag{6.21}$$

It follows that, in the absence of shot noise, the covariance between estimators for the the line correlation function on scales r_i and r_j can be written (Eggemeier and Smith, 2017)

$$\mathrm{Cov}_{\mathrm{G}}\left[\ell(r_i),\ell(r_j)\right] = \frac{(r_i r_j)^{9/2}}{V^3} \iint\limits_{\substack{|\boldsymbol{k}_1|,|\boldsymbol{k}_2|,\\ |\boldsymbol{k}_1+\boldsymbol{k}_2|\leq 2\pi/r}} \frac{\mathrm{d}^3k_1}{k_{\mathrm{f}}^3}\frac{\mathrm{d}^3k_2}{k_{\mathrm{f}}^3}\Big(j_0(|2\boldsymbol{k}_1+\boldsymbol{k}_2|\,r_i)\big[2\,j_0(|\boldsymbol{k}_1-\boldsymbol{k}_2|\,r_j) + j_0(|2\boldsymbol{k}_1+\boldsymbol{k}_2|\,r_j)\big] + r_i\leftrightarrow r_j\Big), \tag{6.22}$$

where $k_{\mathrm{f}} = 2\pi/V^{1/3}$ denotes the fundamental frequency (defined above Eq. (6.9)), and $r = \max\{r_i, r_j\}$. Note that (6.22) is not diagonal; the integral that defines the line correlation function depends on a range of Fourier modes for any scale r_i, and any Fourier modes that are common between $\ell(r_i)$ and $\ell(r_j)$ will contribute a nonzero covariance. Moreover, Eq. (6.22) shows that the Gaussian covariance is independent of redshift and all cosmological parameters.

6.3.4 Modal bispectrum

It was explained in Sec. 6.2.3 that the modal decomposition is defined by choice of a basis Q_n that samples groups of relevant Fourier configurations. The structure and ordering of the Q_n determine those configurations we wish to prioritise. But unless we carefully adjust the Q_n they will be correlated, and these correlations will be inherited by the β_n^Q. The outcome is that the covariance matrix for estimators of the β_n^Q is rather complex.

Construction of R-basis

To avoid this we redefine the basis, as in Eq. (6.8), to simplify the covariance matrix for estimators of the corresponding β_n^R. The construction proceeds in stages. First, consider the expected signal-to-noise with which it is possible to measure a single mode Q_n/w from (6.7). Using a Gaussian approximation for the noise this can be written

$$6\left(\frac{\mathscr{S}}{\mathscr{N}}\right)^2_{Q_n} = \int \frac{\mathrm{d}^3k_1}{(2\pi)^3}\frac{\mathrm{d}^3k_2}{(2\pi)^3}\frac{\mathrm{d}^3k_3}{(2\pi)^3}(2\pi)^3\frac{\delta_{\mathrm{D}}(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)}{w(k_1,k_2,k_3)^2}\frac{Q_n(k_1,k_2,k_3)^2}{P(k_1)P(k_2)P(k_3)}. \tag{6.23}$$

We are free to choose the weight w to simplify this integral. We define[3]

$$w(k_1,k_2,k_3) = \sqrt{\frac{k_1 k_2 k_3}{P(k_1)P(k_2)P(k_3)}}, \tag{6.24}$$

after which the computation of the expected signal-to-noise reduces to

$$6\left(\frac{\mathscr{S}}{\mathscr{N}}\right)^2_{Q_n} = \langle\!\langle Q_n | Q_n \rangle\!\rangle. \tag{6.25}$$

To write this and similar expressions economically we have introduced the notation

$$\langle\!\langle f|g\rangle\!\rangle \equiv \int \frac{\mathrm{d}^3k_1}{(2\pi)^3}\frac{\mathrm{d}^3k_2}{(2\pi)^3}\frac{\mathrm{d}^3k_3}{(2\pi)^3}(2\pi)^3\delta_{\mathrm{D}}(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)\frac{f(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)g(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)}{k_1k_2k_3} \tag{6.26}$$

for any f and g. In the special case that these depend only on the wave numbers k_i and not their orientations $\hat{\boldsymbol{k}}_i$ some of the angular integrations are trivial and we obtain the simpler expression

$$\langle\!\langle f|g\rangle\!\rangle \equiv \frac{1}{8\pi^4}\int_{\mathscr{V}} \mathrm{d}k_1\,\mathrm{d}k_2\,\mathrm{d}k_3\, f(k_1,k_2,k_3)g(k_1,k_2,k_3). \tag{6.27}$$

Here, $\mathscr{V}$ represents the set of points (k_1,k_2,k_3) where lines of length k_1, k_2 and k_3 can be arranged to form a triangle, i.e. $2\max\{k_i\} \le \sum_i k_i$; for details, see Fergusson et al. (2010). In principle the integral can be carried over all k_i, but in practice it will be cut off at upper and lower limits $k_{\max}$ and $k_{\min}$. The expressions (6.26) and (6.27) can be regarded as an inner product on the Q_n that weights each contributing Fourier configuration according to its individual signal-to-noise.

Second, the R-basis is chosen to be diagonal with respect to this inner product. As we will see below, because the resulting R_n modes are orthogonal when weighted by signal-to-noise, the covariance matrix for estimators of the coefficients β_n^R becomes diagonal under the same approximation of Gaussian noise used to determine the weighting in (6.23). Specifically, we define

$$\langle\!\langle Q_m|Q_n\rangle\!\rangle \equiv \gamma_{mn} \equiv \frac{(k_{\max}-k_{\min})^3}{8\pi^4}\bar{\gamma}_{mn}. \tag{6.28}$$

It is sometimes preferable to express results in terms of $\bar{\gamma}_{mn}$, which is independent of $k_{\min}$ and $k_{\max}$. For any suitable Q-basis both γ_{mn} and $\bar{\gamma}_{mn}$ will be symmetric and positive-definite and may be factored into the product of a matrix and its transpose. Therefore there exists a matrix λ_{mn} such that $\bar{\gamma}_{mn} = \sum_r \lambda_{mr}\lambda_{nr}$. Application of (6.8) with λ_{mn} as the transformation matrix yields $R_n = \sum_{n'} \lambda^{-1}_{nn'} Q_{n'}$, and these modes are

[3] In fact for our estimation procedure we shall use the measured power spectrum $\hat{P}$ in place of P in the formula for the weighting function. This makes little difference to the expectation value, but is preferred as it makes no knowledge about the underlying cosmology.

orthogonal in the sense

$$\langle\!\langle R_m | R_n \rangle\!\rangle = \frac{(k_{\max} - k_{\min})^3}{8\pi^4} \mathbf{1}_{mn}. \tag{6.29}$$

Determination of modal coefficients

Whether we work with the Q- or R-basis, we must predict the corresponding β-coefficients for each model of interest. In practice the extra matrix operations needed to obtain the R-basis mean that it is simplest to perform calculations in the Q-basis, before translating to the R-basis to interpret the results. We adopt this procedure whenever concrete calculations using the modal decomposition are required. We use the Q-basis constructed by Fergusson et al. (2010). (The details are summarised in App. 6.A.1.) It is not intended to prioritise any single class of Fourier configurations, but rather attempts to provide a good description of reasonably smooth bispectra over a range of shapes and scales.

To extract the β_n^Q we use (6.26). Assuming (6.7) can be interpreted as an equality, we conclude that for an arbitrary bispectrum $B^{\text{theory}}(k_1, k_2, k_3)$

$$\langle\!\langle wB^{\text{theory}} | Q_m \rangle\!\rangle = \sum_{n=0}^{n_{\max}-1} \beta_n^{Q,\text{theory}} \gamma_{nm}. \tag{6.30}$$

Finally, the individual β_n^Q should be extracted by contraction with the inverse matrix γ_{mn}^{-1}. If the bispectrum has no angular dependence then the inner product can be computed using the simplified expression (6.27), which yields

$$\beta_n^{Q,\text{theory}} = \frac{1}{8\pi^4} \sum_m \gamma_{nm}^{-1} \int_{\mathcal{V}} \mathrm{d}k_1\, \mathrm{d}k_2\, \mathrm{d}k_3\, \sqrt{k_1 k_2 k_3} B_\epsilon^{\text{theory}}(k_1, k_2, k_3) Q_m(k_1, k_2, k_3), \tag{6.31}$$

where we have used the quantity B_ϵ defined in (6.16). The $\beta_n^{R,\text{theory}}$ may be obtained by the transformation $\beta_n^R = \sum_m \lambda_{mn} \beta_m^Q$. The appearance of the phase bispectrum in (6.31) is a consequence of our choice of weight w.

Eq. (6.30) would continue to apply were we to change the definition of the 'inner product' $\langle\!\langle \cdot | \cdot \rangle\!\rangle$, and an analogue of (6.31) would continue to give the individual $\beta_n^{Q,\text{theory}}$. Our choice of signal-to-noise weighting in $\langle\!\langle \cdot | \cdot \rangle\!\rangle$ is important only for construction of the R-modes and the covariance inherited by the $\beta_n^{R,\text{theory}}$.

Numerical evaluation

In practice, Eq. (6.31) requires evaluation of a 3-dimensional integral over the region $\mathcal{V}$. To implement it we compute wB on a 200^3 cubic lattice in (k_1, k_2, k_3) and estimate the integral by volume-weighted cubature over this lattice. Some work is required to account for irregular boundary orientations; we give these details in App. 6.A.2.

Covariance

Finally we compute the covariance of estimators for the β_n^R coefficients under the assumption of Gaussian covariance for the bispectrum estimator $\delta(\boldsymbol{k}_1)\delta(\boldsymbol{k}_2)\delta(\boldsymbol{k}_3)V^{-1} \times \mathbf{1}_{\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3,0}$. Using Eq. (6.26), and (6.30) with R exchanged for Q, we obtain

$$\begin{aligned}\langle \beta_m^R \beta_n^R \rangle = (2\pi)^3\delta(0)\frac{6}{V^2}\frac{(8\pi^4)^2}{(k_{\max}-k_{\min})^6}\int \frac{\mathrm{d}^3k_1\,\mathrm{d}^3k_2\,\mathrm{d}^3k_3}{(2\pi)^9}(2\pi)^3\delta(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3) \\ \times \frac{R_m(k_1,k_2,k_3)R_n(k_1,k_2,k_3)}{k_1k_2k_3} \\ = \frac{6}{V}\frac{(8\pi^4)^2}{(k_{\max}-k_{\min})^6}\langle\!\langle R_m|R_n\rangle\!\rangle. \end{aligned} \tag{6.32}$$

The weighting for each Fourier configuration matches the signal-to-noise, making this correlator diagonal as a consequence of our construction of the R-basis. Therefore we conclude

$$\mathrm{Cov}_{\mathrm{G}}\left(\beta_m^R, \beta_n^R\right) = \frac{6}{V}\frac{8\pi^4}{(k_{\max}-k_{\min})^3}\mathbf{1}_{mn}. \tag{6.33}$$

As for the line correlation function, it is independent of redshift and cosmological parameters. If we were to abandon the approximation of Gaussian covariance then (6.32) would no longer be proportional to exactly $\langle\!\langle R_m|R_n\rangle\!\rangle$. In this case the amplitude of the diagonal elements would be modified, and non-diagonal components would appear.

6.3.5 Galaxy bias

The discussion in Sections 6.3.2–6.3.4 was framed in terms of the dark matter over-density δ, but this is not what is measured by surveys of large-scale structure. Instead, they record the abundance of galaxies or some other population of tracers whose density responds to the dark matter density but need not match it.

On large scales the relation between the galaxy (δ_g) and dark matter (δ) density fields is well-described by the linear model $\delta_g = b_1\delta$ (Fry et al., 1993; Kaiser, 1984). The *linear bias parameter* b_1 may be redshift-dependent, and varies between different populations of galaxies. On small scales the over-densities are larger, and both non-linear and non-local corrections become important. To obtain a satisfactory description we must typically include terms at least quadratic (or higher) in δ (Fry et al., 1993; Smith et al., 2007), together with terms involving the tidal gravitational field (Baldauf et al., 2012; Catelan et al., 2000; Chan et al., 2012; McDonald and Roy, 2009).

In what follows we assume the local Lagrangian bias model, in which the galaxy over-density at early times is taken to be a local function of the dark matter over-density. At later times the bias is determined by propagating this relationship along the dark matter flow. McDonald and Roy (2009) demonstrated that this implies the Eulerian

galaxy over-density at the time of observation can be written

$$\delta_g(\boldsymbol{x}) = b_1\delta(\boldsymbol{x}) + \frac{1}{2}b_2\left[\delta^2(\boldsymbol{x}) - \langle\delta^2(\boldsymbol{x})\rangle\right] + \frac{1}{2}b_{s^2}\left[s^2(\boldsymbol{x}) - \langle s^2(\boldsymbol{x})\rangle\right] + \cdots, \tag{6.34}$$

where '$\cdots$' denotes terms of third order and higher that we have not written explicitly. The field $s^2(\boldsymbol{x}) = s^{ij}(\boldsymbol{x})\,s_{ji}(\boldsymbol{x})$ is a contraction of the tidal tensor, defined by $s_{ij}(\boldsymbol{x}) \equiv \left[\partial_i\partial_j\nabla^{-2} - \frac{1}{3}\mathbf{1}_{ij}\right]\delta(\boldsymbol{x})$. Therefore, up to second order in δ, we require two additional redshift- and population-dependent bias parameters: the *quadratic bias* b_2, as well as the *non-local bias* b_{s^2}. In the local Lagrangian model the non-local bias satisfies $b_{s^2} = -4(b_1-1)/7$ (Baldauf et al., 2012; Chan et al., 2012), although in more general biasing prescriptions it could be allowed to vary independently.

Power spectrum

After translating to Fourier space it follows that the tree-level galaxy power spectrum can be written

$$P^{\text{tree}}_{\text{gal}}(k) = b_1^2 P^{\text{tree}}(k). \tag{6.35}$$

To obtain a consistent result at one-loop we should include the unwritten third-order contributions in (6.34), which generate multiplicative renormalisations of the linear power spectrum in the same way as the '13' terms of one-loop SPT. McDonald and Roy (2009) showed that these could be collected into a single new parameter which we denote $b_{3\text{nl}}$ to match Gil-Marín et al. (2015). Therefore

$$\begin{aligned} P^{\text{1-loop}}_{\text{gal}}(k) = {} & b_1^2 P^{\text{1-loop}}(k) + 2b_1 b_2 P_{b2}(k) + 2b_1 b_{s^2} P_{bs2}(k) + b_2^2 P_{b22}(k) \\ & + 2b_2 b_{s^2} P_{b2,bs2}(k) + b_{s^2}^2 P_{bs22}(k) + 2b_1 b_{3\text{nl}}\sigma_3^2(k) P^{\text{tree}}(k). \end{aligned} \tag{6.36}$$

Saito et al. (2014) showed that in the local Lagrangian model $b_{3\text{nl}}$ satisfies $b_{3\text{nl}} = 32(b_1-1)/315$. Explicit expressions for all terms appearing in (6.36) were given by McDonald and Roy (2009). Note that contributions from the non-linear bias appear only in the one-loop power spectrum. Finally, in the halo model approach we use a simple prescription that replaces $P^{\text{1-loop}}(k)$ in (6.36) with $P^{\text{halo}}(k)$ and leaves the remaining terms unchanged.

Bispectrum

In contrast to the power spectrum, the bispectrum receives corrections from non-linear bias terms even at tree-level. Specifically,

$$B^{\text{tree}}_{\text{gal}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) = b_1^3 B^{\text{tree}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) + b_1^2 P^{\text{tree}}(k_1) P^{\text{tree}}(k_2)\left[b_2 + b_{s^2} S_2(\boldsymbol{k}_1, \boldsymbol{k}_2)\right] + \text{cyclic}, \tag{6.37}$$

where $S_2(\boldsymbol{k}_1, \boldsymbol{k}_2) \equiv (\boldsymbol{k}_1 \cdot \boldsymbol{k}_2)^2/(k_1 k_2)^2 - 1/3$ is the kernel appearing in the Fourier transform of the contracted tidal field, $s^2(\boldsymbol{k}) = (2\pi)^{-3} \int \mathrm{d}^3 q\, S_2(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q})\delta(\boldsymbol{q})\delta(\boldsymbol{k}-\boldsymbol{q})$.

To obtain the galaxy bispectrum consistently at one loop one should compute the dark matter over-density to fourth order in perturbation theory and develop the bias expansion to the same order, which leads to additional bias parameters, as we describe in Chap. 7. However, for reasons of simplicity we here work with an estimate of the one-loop bispectrum by making the approximation

$$\begin{aligned} B_{\mathrm{gal}}^{\text{1-loop}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) = b_1^3 B^{\text{1-loop}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) \\ + b_1^2 P^{\text{1-loop}}(k_1) P^{\text{1-loop}}(k_2) \left[b_2 + b_{s^2} S_2(\boldsymbol{k}_1, \boldsymbol{k}_2)\right] + \text{cyclic}, \end{aligned} \tag{6.38}$$

and in line with the procedure adopted for the power spectrum, we compute the halo model galaxy bispectrum by making the replacements $P^{\text{1-loop}} \to P^{\text{halo}}$ and $B^{\text{1-loop}} \to B^{\text{halo}}$. This approximation is consistent with the prescriptions used by Gil-Marín et al. (2015) and Baldauf et al. (2016), but as it does not account for marginalisation over the extra third and fourth order bias parameters, it has the drawback of potentially overestimating the constraining power of the bispectrum.

Application to bispectrum proxies

The outcome of this discussion is that, to predict the integrated bispectrum, line correlation function, or modal bispectrum for the galaxy density field, we should make the replacements $P^{\text{theory}}(k) \to P_{\mathrm{gal}}^{\text{theory}}(k)$ and $B^{\text{theory}}(k_1, k_2, k_3) \to B_{\mathrm{gal}}^{\text{theory}}(k_1, k_2, k_3)$ where necessary in Eqs. (6.12), (6.15) and (6.31).

To obtain theory predictions at tree-level we use Eqs. (6.35) and (6.37), whereas to obtain predictions at one-loop we use Eqs. (6.36) and (6.38). Finally, to evaluate predictions using the halo model we again apply the latter two expressions, but with the changes indicated above.

6.4 Estimating bispectrum proxies from N-body simulations

In this section we briefly describe our N-body simulations and explain how they are used to estimate the Fourier bispectrum and its proxies ib, ℓ and β_n^Q.

6.4.1 Simulations

Our measurements are based on two sets of simulations: (1) 200 N-body simulations containing dark matter only, with a fixed choice of fiducial cosmological parameters; (2) a total of 60 simulations constructed by varying one cosmological parameter at a

Table 6.1 Fiducial values of the cosmological parameters, together with the step size $\Delta\theta$ used to vary each parameter in the simulations. We perform one simulation with offset $+\Delta\theta$ and one with increment $-\Delta\theta$, giving two offset simulations per parameter. With seven parameters and four realisations per model this gives $4+2\times7\times4=60$ simulations in the suite. The bias parameters are assumed to be $b_1=1$ and $b_2=0$.

Parameter θ	Ω_m	Ω_b	w_0	w_a	σ_8	n_s	h
Fiducial value	0.25	0.040	−1.0	0.0	0.8	1.00	0.70
$\Delta\theta$	±0.05	±0.005	±0.2	±0.1	±0.1	±0.05	±0.05

time, with four realisations per model including the fiducial set. These simulations were performed on the ZBOX supercomputer at the University of Zurich and were described in Smith (2009) and Smith et al. (2014). Each set uses a comoving boxsize of $L=1500\,h^{-1}$Mpc and contains $N=750^3$ particles. Initial conditions for the particles were set at redshift $z=49$ using second-order Lagrangian perturbation theory acting on a realisation of a Gaussian random field (Crocce et al., 2006) with transfer functions from CMBFAST (Seljak and Zaldarriaga, 1996). The particles are evolved to $z=0$ under the influence of gravity using the Gadget-2 code (Springel, 2005), modified to allow a time-evolving equation of state for dark energy.

The fiducial cosmological parameters correspond to a flat ΛCDM model and are summarised in Table 6.1. Specifically, Ω_m and Ω_b are the matter and baryon density parameters; w_0 and w_a parametrise the equation of state for dark energy, viz. (Chevallier and Polarski, 2001; Efstathiou, 1999) $w(a)\equiv w_0+(1-a)\,w_a$; σ_8 is the amplitude of density fluctuations smoothed on a scale $8\,h^{-1}$Mpc; n_s is the spectral index of the primordial power spectrum; and h is the dimensionless Hubble parameter. We collectively write these as a vector θ_α with index α labelling the different parameters. To construct set (2) each parameter is offset by $+\Delta\theta_\alpha$ and $-\Delta\theta_\alpha$, with all other parameters held fixed. The step sizes $\Delta\theta_\alpha$ are listed in Table 6.1. To reduce noise when estimating parameter derivatives, we construct initial conditions for each of the four realisations using the same Gaussian random field as its fiducial partner. Since we vary over seven cosmological parameters this gives a total of $4+2\times7\times4=60$ simulations in the suite.

6.4.2 Density field

To compute the over-density field in each simulation we use the cloud-in-cell assignment scheme to distribute particles over a regular Cartesian grid. We apply a fast Fourier transform and extract the discrete real-space density field by deconvolving the cloud-in-cell window function. The result is

$$\delta^{\rm disc}(\boldsymbol{k})=\frac{\delta^{\rm grid}(\boldsymbol{k})}{W_{\rm CIC}(\boldsymbol{k})},\quad\text{where}\quad W_{\rm CIC}(\boldsymbol{k})=\prod_{i=1}^{3}\left[\frac{\sin\left(\pi k_i/2k_{\rm Ny}\right)}{\pi k_i/2k_{\rm Ny}}\right]^2. \tag{6.39}$$

The labels 'disc' and 'grid' label Fourier-space fields in the full volume V and on the cloud-in-cell grid, respectively. The Nyquist frequency $k_{\rm Ny} = \pi N_{\rm grid}/L$ is determined by the number of grid cells per dimension. For our numerical results we use $N_{\rm grid} = 512$.

6.4.3 Estimating the power spectrum

Given a realisation of the δ-field within a simulation volume $V = L^3 = (2\pi)^3\delta_{\rm D}(0)$, a simple estimator for the power at wave vector $\boldsymbol{k}_1$ can be written[4]

$$\hat{\mathscr{P}}(\boldsymbol{k}_1, \boldsymbol{k}_2) = \delta(\boldsymbol{k}_1)\delta(\boldsymbol{k}_2)\mathbf{1}_{\boldsymbol{k}_1,-\boldsymbol{k}_2}/V\,. \tag{6.40}$$

Unfortunately this procedure is very noisy. An improved estimate can be obtained by summing over a set of modes satisfying the closure criterion $\sum_i \boldsymbol{k}_i = 0$ within a thin $\boldsymbol{k}$-shell. Since we are working in finite volume the available modes are discretised in units of the fundamental frequency $k_{\rm f} = 2\pi/L$, and therefore the thin-shell average should be written

$$\hat{P}(k) = \frac{1}{V_P(k)}\int \mathrm{d}^3q_1\,\mathrm{d}^3q_2\,\delta_{\rm D}(\boldsymbol{q}_1+\boldsymbol{q}_2)\hat{\mathscr{P}}(\boldsymbol{q}_1,\boldsymbol{q}_2)\tilde{\Pi}_k(\boldsymbol{q}_1)\tilde{\Pi}_k(\boldsymbol{q}_2)\,, \tag{6.41}$$

where $\Delta k \geq k_{\rm f}$ represents a bin width, and we have introduced the binning function $\tilde{\Pi}_k(\boldsymbol{q})$ which is defined to be unity if $|\boldsymbol{q}| \in [k-\Delta k/2, k+\Delta k/2]$ and zero otherwise. Finally, the quantity V_P represents the volume of the spherical shell accounting for discretisation,

$$\begin{aligned} V_P(k) &\equiv \int \mathrm{d}^3q_1\,\mathrm{d}^3q_2\,\delta_{\rm D}(\boldsymbol{q}_1+\boldsymbol{q}_2)\tilde{\Pi}_k(\boldsymbol{q}_1)\tilde{\Pi}_k(\boldsymbol{q}_2) = \int \mathrm{d}^3q\,\tilde{\Pi}_k^2(\boldsymbol{q}) = \int \mathrm{d}^3q\,\tilde{\Pi}_k(\boldsymbol{q}) \\ &= 4\pi k^2\Delta k\Big[1+\frac{1}{12}\Big(\frac{\Delta k}{k}\Big)^2\Big]\,. \end{aligned} \tag{6.42}$$

6.4.4 Estimating the bispectrum

In analogy with the power spectrum, an estimator for a single configuration of the Fourier bispectrum can be written $\hat{\mathscr{B}}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) = \delta(\boldsymbol{k}_1)\delta(\boldsymbol{k}_2)\delta(\boldsymbol{k}_3)\mathbf{1}_{\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3,0}/V$. [This expression was already used in Sec. 6.3.4 to obtain the Gaussian covariance for estimators of the β_n^R.] To obtain an acceptable signal-to-noise we should again average over a set of configurations whose wave numbers lie within suitable discretised $\boldsymbol{k}$-shells. After doing so we obtain the estimator

$$\begin{aligned} \hat{B}(k_1,k_2,k_3) = \frac{1}{V_B(k_1,k_2,k_3)}\int \mathrm{d}^3q_1\,\mathrm{d}^3q_2\,\mathrm{d}^3q_3\,\delta_{\rm D}(\boldsymbol{q}_1+\boldsymbol{q}_2+\boldsymbol{q}_3)\,\hat{\mathscr{B}}(\boldsymbol{q}_1,\boldsymbol{q}_2,\boldsymbol{q}_3) \\ \times\,\tilde{\Pi}_{k_1}(\boldsymbol{q}_1)\,\tilde{\Pi}_{k_2}(\boldsymbol{q}_2)\,\tilde{\Pi}_{k_3}(\boldsymbol{q}_3)\,, \end{aligned} \tag{6.43}$$

[4] In the remainder of this chapter we assume it is understood that we are dealing with the discrete density field whenever we refer to measured quantities, and drop the label 'disc'.

where the normalisation V_B should now be evaluated using (Joachimi et al., 2009; Sefusatti et al., 2006)

$$V_B(k_1,k_2,k_3) \equiv \int \mathrm{d}^3q_1\,\mathrm{d}^3q_2\,\mathrm{d}^3q_3\,\delta_\mathrm{D}(\boldsymbol{q}_1+\boldsymbol{q}_2+\boldsymbol{q}_3)\,\tilde{\Pi}_{k_1}(\boldsymbol{q}_1)\,\tilde{\Pi}_{k_2}(\boldsymbol{q}_2)\,\tilde{\Pi}_{k_3}(\boldsymbol{q}_3)$$
$$\approx 8\pi^2 k_1 k_2 k_3\,(\Delta k)^3\,. \tag{6.44}$$

Dividing by the square of the fundamental cell volume shows that the number of configurations scales as $N_\mathrm{triangles}(k_1,k_2,k_3) = V_B(k_1,k_2,k_3)/k_\mathrm{f}^6 \propto N_1N_2N_3$, where $N_i \equiv k_i/k_\mathrm{f}$ is the length of the side k_i in units of the fundamental mode. Hence, if we scale the configuration by $k_i \to \lambda k_i$ then the number of available configurations scales as λ^3.

Sefusatti (2005), Fergusson, Regan, and Shellard (2012) and Scoccimarro (2015) observed that (6.43) could be implemented efficiently by rewriting the Dirac δ-function using its Fourier representation, $(2\pi)^3\delta_\mathrm{D}(\boldsymbol{q}) = \int \mathrm{d}^3x\,\mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{x}}$, and factorising the dependence on the $\boldsymbol{q}_i$. This yields

$$\hat{B}(k_1,k_2,k_3) = \frac{k_\mathrm{f}^3}{(2\pi)^6 V_B(k_1,k_2,k_3)} \int \mathrm{d}^3x\,\mathscr{D}_{k_1}(\boldsymbol{x})\mathscr{D}_{k_2}(\boldsymbol{x})\mathscr{D}_{k_3}(\boldsymbol{x})\,, \tag{6.45}$$

where

$$\mathscr{D}_k(\boldsymbol{x}) \equiv \int \mathrm{d}^3q\,\mathrm{e}^{\mathrm{i}\boldsymbol{x}\cdot\boldsymbol{q}}\delta(\boldsymbol{q})\tilde{\Pi}_k(\boldsymbol{q})\,. \tag{6.46}$$

Similarly,

$$V_B(k_1,k_2,k_3) = \int \frac{\mathrm{d}^3x}{(2\pi)^3}\Pi_{k_1}(\boldsymbol{x})\Pi_{k_2}(\boldsymbol{x})\Pi_{k_3}(\boldsymbol{x}), \tag{6.47}$$

where $\Pi_k(\boldsymbol{x})$ is the inverse Fourier transform of $\tilde{\Pi}_k(\boldsymbol{q})$.

Eq. (6.45) is numerically more efficient than a direct implementation of (6.43), because it requires only three Fourier transforms to compute $\mathscr{D}_k$ for each wave number in the triplet $\{k_1,k_2,k_3\}$. Moreover, once each $\mathscr{D}_k$ has been obtained it can be re-used for any configuration that shares the same wave number. In spite of this improvement, however, it remains a formidable computational challenge to estimate all bispectrum configurations contained within a large volume V. Different strategies have been employed to make the calculation feasible. One option is to coarsely bin configurations with binning width equal to several times the fundamental mode. This drastically reduces the number of configurations to be measured. An alternative is to search only among a limited subset of configurations. This may be helpful if we wish to search for specific physical effects, but risks overlooking important signals if we are searching blindly. In either case the analysis is unlikely to be optimal because information is lost.

6.4.5 Estimating the integrated bispectrum

Our procedure to estimate the integrated bispectrum is based directly on its definition. We separate the total volume into N_s subvolumes, enumerated by the labels $i = 1, \ldots, N_s$. We compute the mean over-density $\hat{\bar{\delta}}_i$ and power spectrum $\hat{P}(k)_i$ within each subvolume. Finally, we average the product $\hat{P}(k)_i \hat{\bar{\delta}}_i$ over all subvolumes. Therefore,

$$\widehat{iB}(k) = \frac{1}{N_s} \sum_{i=1}^{N_s} \hat{P}(k)_i \hat{\bar{\delta}}_i \,. \tag{6.48}$$

The normalised integrated bispectrum can be obtained by rescaling,

$$\widehat{ib}(k) = \frac{\widehat{iB}(k)}{\hat{P}(k)\hat{\sigma}_L^2} \,, \tag{6.49}$$

where here $\hat{P}(k) = \sum_{i=1}^{N_s} \hat{P}(k)_i / N_s$ is the average subvolume power spectrum and $\hat{\sigma}_L^2 = \sum_{i=1}^{N_s} \hat{\bar{\delta}}_i^2 / N_s$ is the average variance of the mean over-density.

6.4.6 Estimating the line correlation function

A procedure to estimate the line correlation function was outlined by Eggemeier and Smith (2017). We evaluate

$$\hat{\ell}(r) = \left(\frac{r^3}{V}\right)^{3/2} \sum_{\substack{|\boldsymbol{k}_1|,|\boldsymbol{k}_2|,\\ |\boldsymbol{k}_1+\boldsymbol{k}_2| \leq 2\pi/r}} \overline{j_0}(|\boldsymbol{k}_1 - \boldsymbol{k}_2| r)\, \epsilon(\boldsymbol{k}_1)\, \epsilon(\boldsymbol{k}_2)\, \epsilon(-\boldsymbol{k}_1 - \boldsymbol{k}_2) \,, \tag{6.50}$$

where $\overline{j_0}(|\boldsymbol{k}|r)$ denotes an average of $j_0(kr)$ taken over the volume of a fundamental k-space cell centred at $\boldsymbol{k}$. The sum scales as $\sim (2L/r)^6$, making its evaluation fast on large scales but challenging on small ones, where the sum includes the majority of Fourier modes. On scales below $\sim 105\,h^{-1}$Mpc we find that the real space estimator described by Eggemeier and Smith (2017) becomes more efficient and therefore we use it within that regime. For scales accessible to both schemes we verified that both estimators yield the same result.

6.4.7 Estimating the modal bispectrum

Eq. (6.31) shows that an estimate of the modal coefficient β_m^Q requires evaluation of $\langle\!\langle w\hat{\mathscr{B}}|Q_n\rangle\!\rangle$, where $\hat{\mathscr{B}}$ is the bispectrum estimator defined in Sec. 6.4.4. Using Eq. (6.26), writing the δ-function using its Fourier representation, and factorising the integral as described in Sec. 6.4.4, we find

$$\langle\!\langle w\hat{\mathscr{B}}|Q_n\rangle\!\rangle = \frac{1}{V} \int \mathrm{d}^3 x\, \mathscr{M}_{n_1}(\boldsymbol{x}) \mathscr{M}_{n_2}(\boldsymbol{x}) \mathscr{M}_{n_3}(\boldsymbol{x}) \,, \tag{6.51}$$

where

$$\mathscr{M}_n(\boldsymbol{x}) \equiv \int \frac{\mathrm{d}^3 k}{(2\pi)^3} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \frac{q_n(k)}{\sqrt{k\hat{P}(k)}} \delta(\boldsymbol{k}) \,. \tag{6.52}$$

Here, $q_n(k)$ is a polynomial used in the construction of the modes Q_n; see App. 6.A.1. Eq. (6.51) shows that the computation can be reduced to a single 3-dimensional integral over the $\mathscr{M}_n(\boldsymbol{x})$, which are themselves weighted Fourier transforms of δ. Finally, β_m^Q can be estimated by contracting with the inverse inner product matrix γ_{mn}^{-1} defined in (6.28),

$$\hat{\beta}_m^Q = \sum_{n=0}^{n_{\max}-1} \langle\!\langle w\hat{\mathscr{B}} | Q_n \rangle\!\rangle \gamma_{nm}^{-1}. \tag{6.53}$$

To obtain the corresponding R-basis coefficients requires a further linear transformation

$$\hat{\beta}_n^R = \sum_m \lambda_{mn} \hat{\beta}_m^Q, \tag{6.54}$$

where λ_{mn} is the matrix defined above (6.29). As explained in Sec. 6.3.4, we generally perform numerical calculations in the Q-basis in order to preserve the simplicity of (6.51), but present results in the R-basis because their covariance properties make these coefficients simpler to interpret. In either basis, the measured coefficients can be used to reconstruct the bispectrum for any required Fourier configuration using Eq. (6.8).

Note that, because the matrix γ_{nm} can be tabulated, measuring a single modal coefficient has the same computational complexity as measuring a single configuration of the Fourier bispectrum.

6.4.8 Choice of bins

In Table 6.2 we summarise the parameters used in implementing estimators for each of these statistical quantities. The power spectrum and Fourier bispectrum are binned by averaging over shells of width Δk as explained in Sections 6.4.3–6.4.4. For the same reasons we also average the subvolume power spectra used to construct the integrated bispectrum. The line correlation function and modal coefficients do not involve averaging over shells, but instead are evaluated using Eqs. (6.50) and (6.51) which are themselves aggregates over groups of configurations. For each statistic we report the minimum and maximum k-modes that contribute, and the total number of measurements or bins. Note that the bispectrum bin width corresponds to $\Delta k = 8\, k_{\mathrm{f}}$.

In what follows we will label the Fourier configurations for the bispectrum using the scheme of Gil-Marín et al. (2017). We assign the label (or 'index') zero to the equilateral configuration with $k_1 = k_2 = k_3 = k_{\min}$. The remaining configurations are ordered so that $k_1 \leq k_2 \leq k_3$ and $k_3 \leq k_1 + k_2$. Their labels are assigned by sequentially increasing k_3, k_2 and k_1 (in this order) and incrementing the index for each valid triangle.

Table 6.2 Shell widths Δk used to average estimators for the power spectrum and bispectrum (where used), together with minimum and maximum modes $k_{\rm min}$, $k_{\rm max}$ and the total number of bins or measurements $N_{\rm bin}$.

	Δk [hMpc^{-1}]	$k_{\rm min}$ [hMpc^{-1}]	$k_{\rm max}$ [hMpc^{-1}]	$N_{\rm bin}$
P	0.010	0.004	0.300	30
ib	0.010	0.021	0.306	29
ℓ	–	0.016	0.314	30
β	–	0.004	0.302	50
B	0.034	0.004	0.302	95

In our measurements of the integrated bispectrum we split the simulation box into 125 subcubes, corresponding to a side of 300 h^{-1}Mpc. This increases $k_{\rm min}$ by a factor of five compared to the full box. Finally, for the line correlation function we use a non-regular r-spacing, spanning the range from 10 to 200 h^{-1}Mpc. The first seven bins are separated by 2.5 h^{-1}Mpc, which doubles to 5 h^{-1}Mpc for the next eleven and to 10 h^{-1}Mpc for the remaining twelve bins.

6.5 Comparison of theoretical predictions and simulations

In this section we present estimates of the typical values for each bispectrum proxy introduced in Sec. 6.2, and implemented using the formulae of Sec. 6.4. We derive these from the 200 simulations of our fiducial cosmology in set (1)—see Sec. 6.4.1—at redshifts $z = 0$, $z = 0.52$ and $z = 1$. Also, using the simulation set (2) we determine how each proxy responds to changes in the cosmological parameters (Sec. 6.5.2). These measurements enable us to characterise the accuracy of the theoretical predictions for these typical values discussed in Sec. 6.3. Finally, in Sec. 6.5.3 we discuss measurements of the covariances and cross-covariances for each pair of proxies.

6.5.1 Mean values in the fiducial cosmology

Comparison of measurements and theoretical predictions

In Figs. 6.1–6.4 we show measurements of each proxy for all three redshifts, averaged over the 200 different realisations. We do not explicitly display our power spectrum measurements, which have been well-studied by previous authors (e.g. Lokas et al., 1996; Makino et al., 1992; Mead et al., 2015; Peacock and Smith, 2000; Scoccimarro and Frieman, 1996; Scoccimarro and Sheth, 2002; Scoccimarro et al., 1998a; Scoccimarro et al., 2001; Seljak, 2000; Smith et al., 2003). In each figure, the top row contrasts our

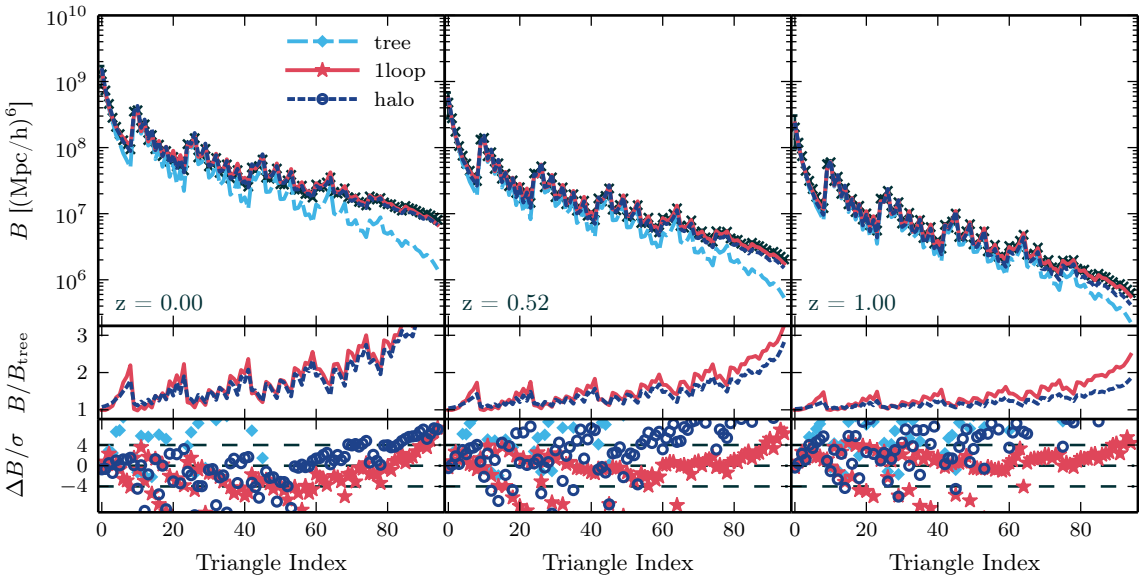


Fig. 6.1 *Top row*: Measurements of the bispectrum as a function of configuration index (see text), estimated from 200 N-body simulations at redshifts $z = 0$, 0.52 and 1. We compare these measurements to the theoretical estimates of Sec. 6.3: the tree-level predictions are shown as dashed light-blue lines, the one-loop predictions are shown as solid red lines, and the halo model predictions are shown as short-dashed dark-blue lines. Black crosses mark the measured values. *Middle row*: One-loop and halo model predictions relative to the tree-level prediction. *Bottom row*: Differences between N-body measurements and theoretical predictions (i.e., $\Delta B = B^{\rm data} - B^{\rm theory}$), normalised to the corresponding 1σ standard deviation in the N-bodyvalue.

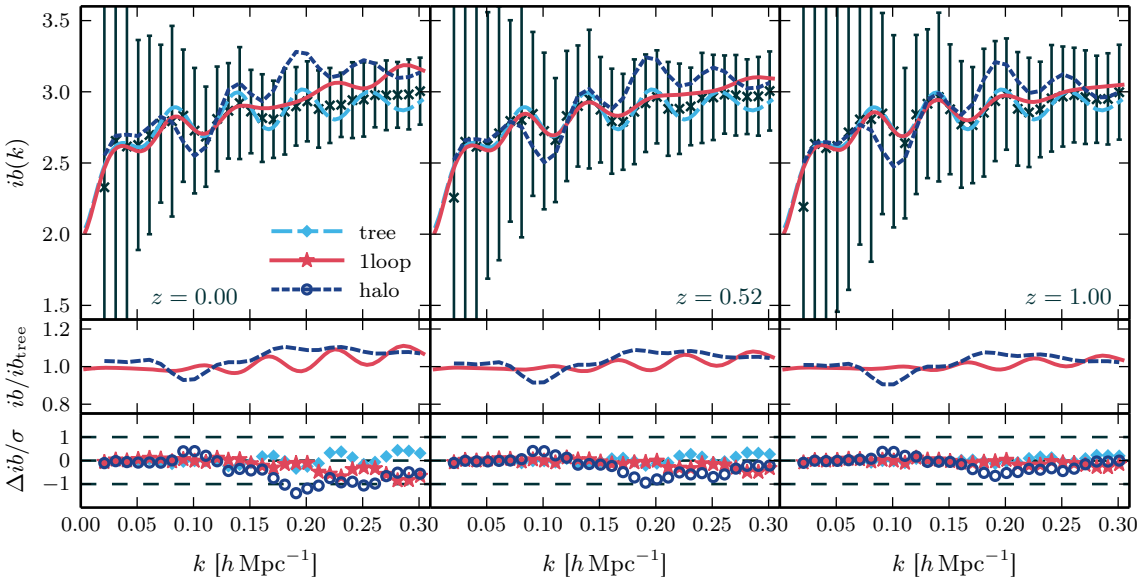


Fig. 6.2 Same configuration as Fig. 6.1, showing values for the normalised integrated bispectrum. Error bars show the 1σ interval.

N-body measurements with the tree-level, one-loop and halo model predictions. The middle row displays the one-loop and halo model predictions relative to the tree-level prediction, and the bottom row shows the difference between the N-body measurements and the theoretical prediction in units of the standard deviation of the N-body estimate.

Fourier bispectrum: We find that both of the SPT predictions are more accurate at large scales and high redshifts. The halo model prediction is a better match at low redshift. The differences between each theoretical estimate and the typical values measured from simulation are broadly consistent with previous analyses; see Lazanu et al. (2016); Schmittfull et al. (2013); Scoccimarro et al. (1998a); Scoccimarro et al. (2001).

Integrated bispectrum: We give values for the normalised integrated bispectrum in Fig. 6.2. Except for a few k-bins the error bars are too large to show any preference for a particular theoretical model. In contrast to Figs. 6.1, 6.3 and 6.4, the bottom row shows that tree-level SPT is a good match to the measured ib at all three redshifts. Conversely, the halo model prediction is a better match at high redshift. Our theoretical predictions are consistent with those reported by Chiang et al. (2014), but our measured values have larger error bars because we work with a smaller simulation volume.

Line correlation function: We present our measurements of the line correlation function in Fig. 6.3. The one-loop and halo-model predictions appearing here are new, and have not previously been studied. The most striking feature is the discrepancy between the halo model and SPT-based predictions in the smallest r-bins. This is consistent with the analyses of Wolstenhulme et al. (2015) and Eggemeier and Smith (2017), which both found differences between the tree-level prediction and values measured from simulation on scales with $r \lesssim 30\,h^{-1}\mathrm{Mpc}$. The agreement is good for larger r.

Modal bispectrum: Finally, in Fig. 6.4 we plot the Fourier bispectrum reconstructed from (6.7) using our measurements of the β_n^Q coefficients. This is easier to interpret than the β-values themselves. The scatter between predicted and measured values (most clearly visible in the bottom row) is similar to the scatter for the directly-measured Fourier bispectrum (Fig. 6.1), and indicates that differences between the reconstructed and directly-measured values are small. We give a more detailed analysis of the accuracy of the modal bispectrum in Sec. 6.5.1.

Theory error: The bottom panels of Figs. 6.1–6.4 show that our theoretical predictions are accurate within a restricted range of scales. Outside this range it becomes progressively more difficult to model the observables. This mis-modelling should be regarded as an additional source of systematic error—a *theory error*—when forecasting constraints, or analysing data, using any of these theoretical models. In particle phenomenology such theory errors are routinely estimated when performing fits to data,

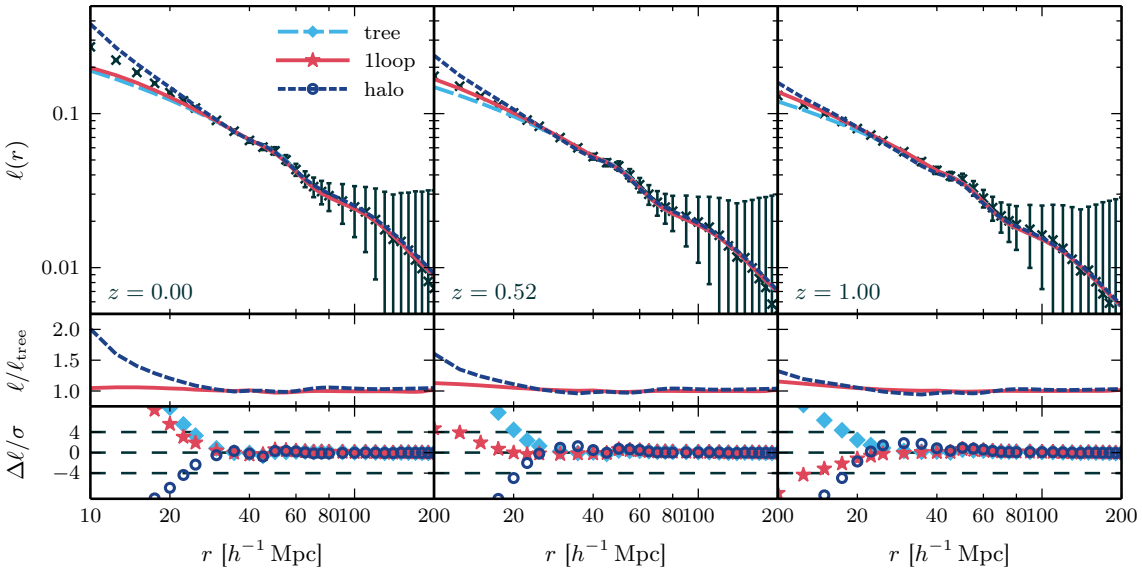


Fig. 6.3 Same configuration as Fig. 6.1, showing values for the line correlation function at scale r. Error bars show the 1σ interval.

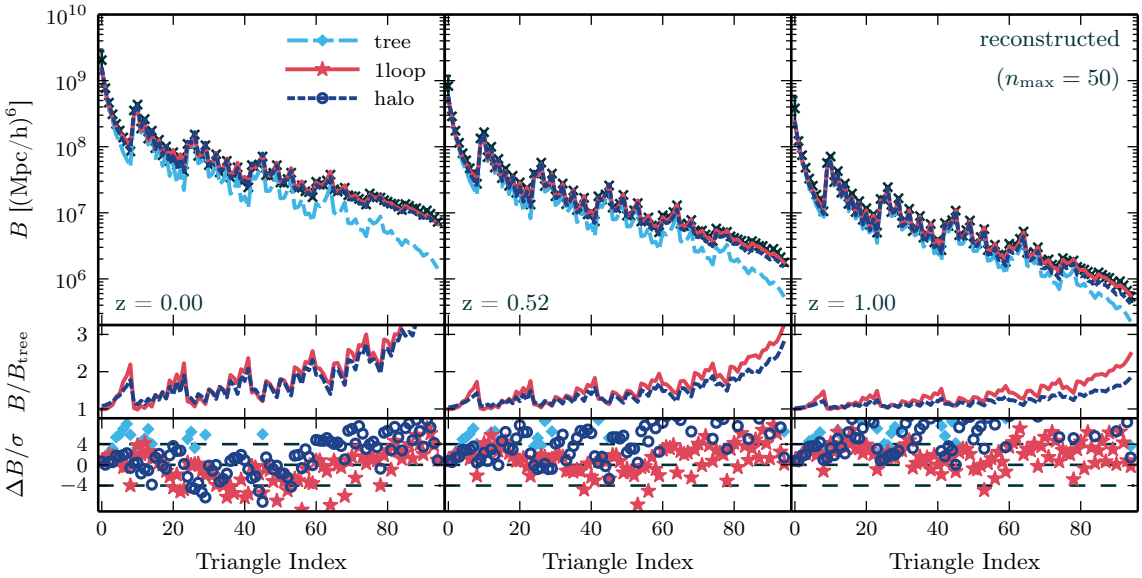


Fig. 6.4 Same configuration as Fig. 6.1, showing values for the Fourier bispectrum reconstructed from the modal coefficients β_n^Q using Eq. (6.7). In the bottom row we plot differences computed using $\Delta B = B_{\rm modal} - B^{\rm theory}$.

but their use in cosmology is less common. In this chapter we construct Fisher forecasts for parameter error bars using both SPT-based models and the halo model. Comparison of these error bars enables us to estimate the impact of theoretical uncertainties on future constraints that incorporate three-point statistics (see Sec. 6.7.4).

An alternative prescription for estimating theory errors was used by Baldauf et al. (2016) and Welling et al. (2016). In their approach the theoretical uncertainty in one-loop SPT is estimated from the next-order term in the loop expansion. We find that this prescription gives noticeably larger estimates than the difference between one-

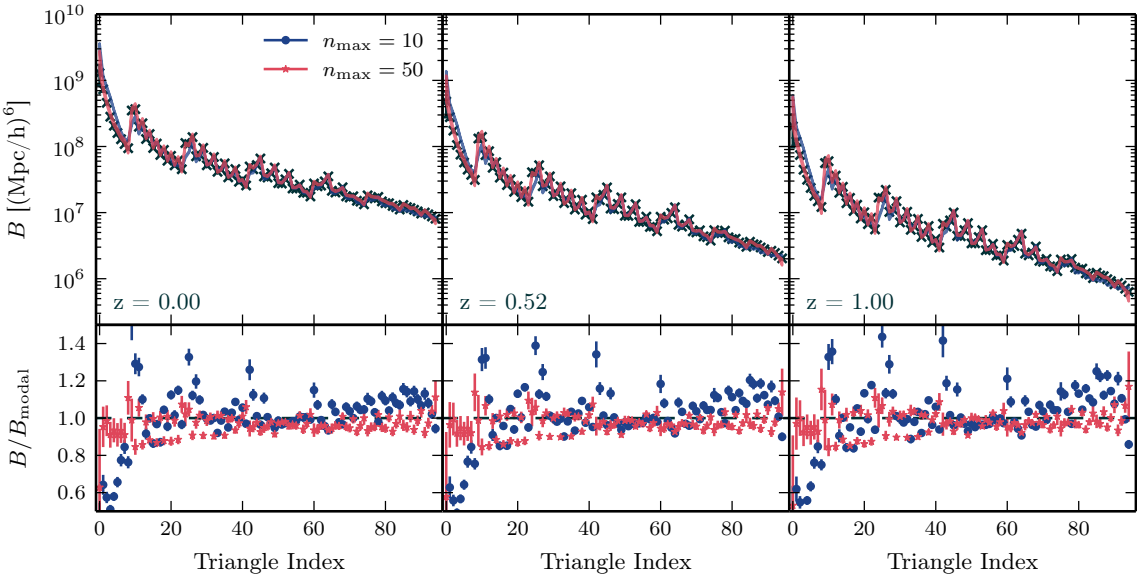


Fig. 6.5 Modal bispectra reconstructed using 10 modes (blue) and 50 modes (red) at redshifts $z = 0$, 0.52 and 1. The lower panels show the ratio of the measured normal bispectrum and modal bispectrum.

loop SPT and the values we measure from simulations. Therefore, although Baldauf et al. (2012) and Welling et al. (2016) concluded that (for example) constraints on some types of primordial non-Gaussianity would be weakened significantly after accounting for theory errors, our numerical comparison suggests that the attainable error may degrade by less than their analysis would suggest. One might object that the differences between our simulation results and the one-loop predictions are fortuitously small for our set of fiducial cosmological parameters. Although this objection is justified, we will demonstrate in Sec. 6.5.2 that there is also a good match between the measured bispectrum derivatives with respect to the cosmological parameters and their one-loop predictions. This indicates that we can expect an equally good fit of the one-loop model for parameters that are at least in the vicinity of the fiducial set.

Accuracy of modal reconstruction

Comparison of Figs. 6.1 and 6.4 demonstrates that the Fourier bispectrum reconstructed from our measurements of the β_n^Q accurately reproduces the correct amplitude and shape dependence. This information is embedded in the modal coefficients. For example, the zeroth basis mode $R_0 \propto Q_0$ is a constant and therefore $\beta_0^R \propto \beta_0^Q$ captures information about the mean amplitude of the Fourier bispectrum over all configurations—or, equivalently, the skewness of δ. The next few modes are slowly varying functions of configuration. Taken together, these low-order modes carry the principal amplitude information and for reasonably smooth bispectra we expect they exhibit the strongest dependence on background cosmological parameters. The higher modes capture more

subtle detail. As with any basis decomposition, their inclusion increases the accuracy of the reconstruction.

To see this in detail, consider a reconstruction using only $n_{\rm max} = 10$ modes. In Fig. 6.5 we plot the Fourier bispectrum reconstructed in this way (blue line) compared to the reconstruction using $n_{\rm max} = 50$ described above (red line). Black crosses mark the measured data points. In the lower panel we plot the ratio between these measured values and the reconstructions. It should be highlighted that the Fourier bispectrum values are aggregations within cubes of side $\Delta k = 8k_{\rm f}$; on large scales (corresponding approximately to small triangle index) one should expect that the bispectrum value might be systematically biased away from the true value. However, for large triangle indices it is clear that the modal estimator faithfully reconstructs the underlying bispectrum. The accuracy is good whether we use $n_{\rm max} = 10$ or $n_{\rm max} = 50$, but the scatter is smaller for $n_{\rm max} = 50$. We conclude that, in this case, the first 10 modes are sufficient to capture the main behaviour of the Fourier bispectrum, but extra modes are helpful if we wish to reproduce the precise configuration dependence to within $\lesssim 10\%$ accuracy.

6.5.2 Derivatives with respect to cosmological parameters

In the remainder of this chapter our aim is to obtain Fisher forecasts of error bars for a parameter set θ_α, where the index α labels one of the cosmological parameters of Table 6.1. For this purpose the role of a theoretical model is to predict the derivatives of observables with respect to each parameter, and the accuracy of the forecast depends on the reliability of these predictions. In this section we study how well our three theoretical models reproduce the derivatives estimated from our simulation suite. We compute the derivative of some estimator $\hat{X}$ at wave number k with respect to a parameter θ_α by the rule

$$\frac{\mathrm{d}\hat{X}(k\,|\,\boldsymbol{\theta})}{\mathrm{d}\theta_\alpha} = \hat{\bar{X}}(k\,|\,\boldsymbol{\theta})\frac{\mathrm{d}\ln\hat{X}(k\,|\,\boldsymbol{\theta})}{\mathrm{d}\theta_\alpha}, \tag{6.55}$$

where $\hat{\bar{X}}(k|\boldsymbol{\theta})$ is the average over the 200 fiducial simulations of set (1) (described in Sec. 6.4.1) for $X \in \{P, B, \beta, ib, \ell\}$, and the logarithmic derivative with respect to θ_α is computed using

$$\frac{\mathrm{d}\ln\hat{X}(k\,|\,\boldsymbol{\theta})}{\mathrm{d}\theta_\alpha} = \frac{1}{4}\sum_{i=1}^{4}\frac{\hat{X}^{(i)}(k\,|\,\boldsymbol{\theta}+\Delta\theta_\alpha) - \hat{X}^{(i)}(k\,|\,\boldsymbol{\theta}-\Delta\theta_\alpha)}{2\Delta\theta_\alpha\hat{X}^{(i)}(k\,|\,\boldsymbol{\theta})}. \tag{6.56}$$

The sum is over the four realisations used in simulation set (2), and the derivative is constructed using the $+\Delta\theta_\alpha$ and $-\Delta\theta_\alpha$ offset simulations described in Sec. 6.4.1. The advantage of the logarithmic derivative is that both realisations in the numerator on the right-hand side of (6.56) share initial conditions with their fiducial partner in

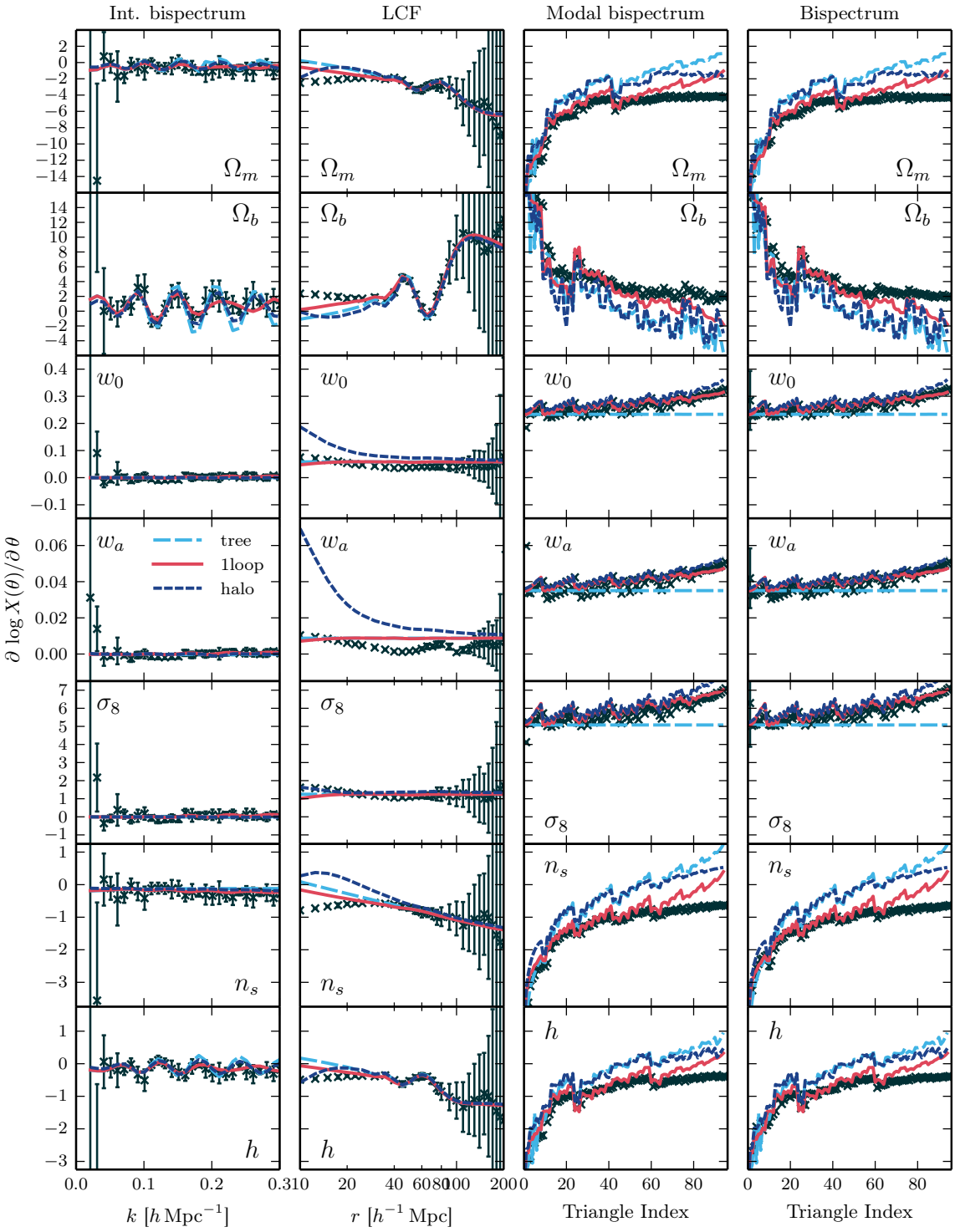


Fig. 6.6 Derivatives of the Fourier bispectrum and its proxies with respect to the parameters at $z = 0$. The four columns show (from left to right) the derivatives of: ib, ℓ, $B_{\rm modal}$ (reconstructed from β) and B. As in Figs. 6.1–6.4, measured values are in black, while light blue dashed, red solid and dark blue short-dashed lines are the tree-level, 1-loop and halo model predictions, respectively.

the denominator. Therefore, division by the fiducial estimate $\hat{X}^{(i)}(k \mid \boldsymbol{\theta})$ minimises dependence on the specific realisation.[5]

In Fig. 6.6 we plot the derivatives of each observable with respect to the cosmological parameters at $z = 0.52$. Our forecasts use three redshift bins, but their behaviour is similar to the $z = 0.52$ bin and the statements made below can be taken to apply at all three redshifts. We do not include the power spectrum, for which the derivatives appeared in Smith et al. (2014).

Integrated bispectrum

The derivatives of the integrated bispectrum are shown in the first column of Fig. 6.6. The error bars on the measured values are too large to show a clear preference for any model—and they are generally so large that the measurement is not significantly different from zero. These results are consistent with those reported by Chiang (2015) for a range of values of Ω_m, σ_8 and n_s. We conclude that the integrated bispectrum is rather insensitive to the background cosmology and is therefore a comparatively poor tool to constrain it. While this means we must expect a Fisher forecast to predict weaker error bars for the parameters of Table 6.1, this insensitivity could be an advantage if the intention is to use the integrated bispectrum as a probe of other physics. For example, in addition to the background cosmology we may wish to use the large-scale structure bispectrum to constrain the possibility of *primordial* three-point configurations produced by inflation on squeezed configurations. Insensitivity to the background cosmology would reduce the likelihood of degeneracies in these measurements.

Line correlation function

The second column of Fig. 6.6 shows the derivatives of the line correlation function. As for the typical values discussed above, the values predicted by our theoretical models are significantly discrepant with the measured values in the smallest r bins. Also, the derivative with respect to the dark energy parameter w_a is particularly discrepant for the halo model. One possible explanation is the construction of the halo model, with its fixed halo mass function and halo profile. Alternatively, it is possible that the halo model power spectrum and bispectrum that we use are subtly inconsistent in a way that produces inaccuracies in the line correlation function on small scales.

[5] This strategy is less successful for the line correlation function. In this case the fiducial value could be very close to zero on some scales. In turn, this produces large errors in the logarithmic derivative. Therefore, for the line correlation function, we estimate the linear derivative $\mathrm{d}\ell/\mathrm{d}\theta_\alpha$ instead.

Modal bispectrum

To simplify comparison of the modal bispectrum with the Fourier bispectrum, Fig. 6.6 plots derivatives of the reconstructed bispectrum rather than derivatives of β_n^Q or β_n^R. Comparison of the final two columns shows that the cosmology-dependence is accurately captured using $n_{\max} = 50$, either for theoretical predictions or the measured values.

There is a significant spread in performance of the theoretical models, with tree-level SPT and the halo model generally offering the poorest match. For the derivatives with respect to Ω_m, Ω_b, n_s and h these models give similar predictions. The probable reason is that, in the standard halo model, as alluded to above, the halo mass function and halo profile are fixed to the fiducial cosmology. Only the input power spectrum is taken to vary with the cosmological parameters, and since it matches the tree-level SPT prediction its derivatives will be equal. Therefore the halo-model derivatives will differ from those of tree-level SPT only via a (possibly scale-dependent) prefactor. More complex halo models with cosmology-dependent halo parametrisations have been studied (see Mead et al., 2016, for an application to dark energy models). However, determining which variation of the halo model captures the cosmological parameter dependence of the bispectrum most accurately is outside the scope of this chapter. We simply note that, if the halo model is to be used for analysis or forecasting of the Fourier bispectrum, its implementation should be chosen with care because its performance depends on these details.

6.5.3 Non-Gaussian covariance

The analytic, Gaussian covariance of each proxy is most accurate at high redshifts and on large scales, where the matter fluctuations are more nearly Gaussian and therefore more accurately described by the power spectrum alone. At low redshifts and on small scales, however, the Gaussian approximation fails due to non-linear evolution of matter fluctuations. This evolution generates additional contributions to the covariance through higher-order n-point correlations.

The simplest and most robust approach to obtain accurate non-Gaussian covariances has been to analyse large suites of N-body simulations. This method was used by Blot et al. (2016); Takahashi et al. (2009); Takahashi et al. (2011), and Klypin and Prada (2018) to study the non-Gaussian covariance of the power spectrum. Other authors have performed analogous studies for the bispectrum (Chan and Blot, 2017; Sefusatti et al., 2006), the real-space partner of the integrated bispectrum (Chiang et al., 2015), and the line correlation function (Eggemeier and Smith, 2017). In this section, we present our measurements of the non-Gaussian covariance for each proxy, estimated from our suite of simulations. We also discuss the cross-covariance between pairs of proxies.

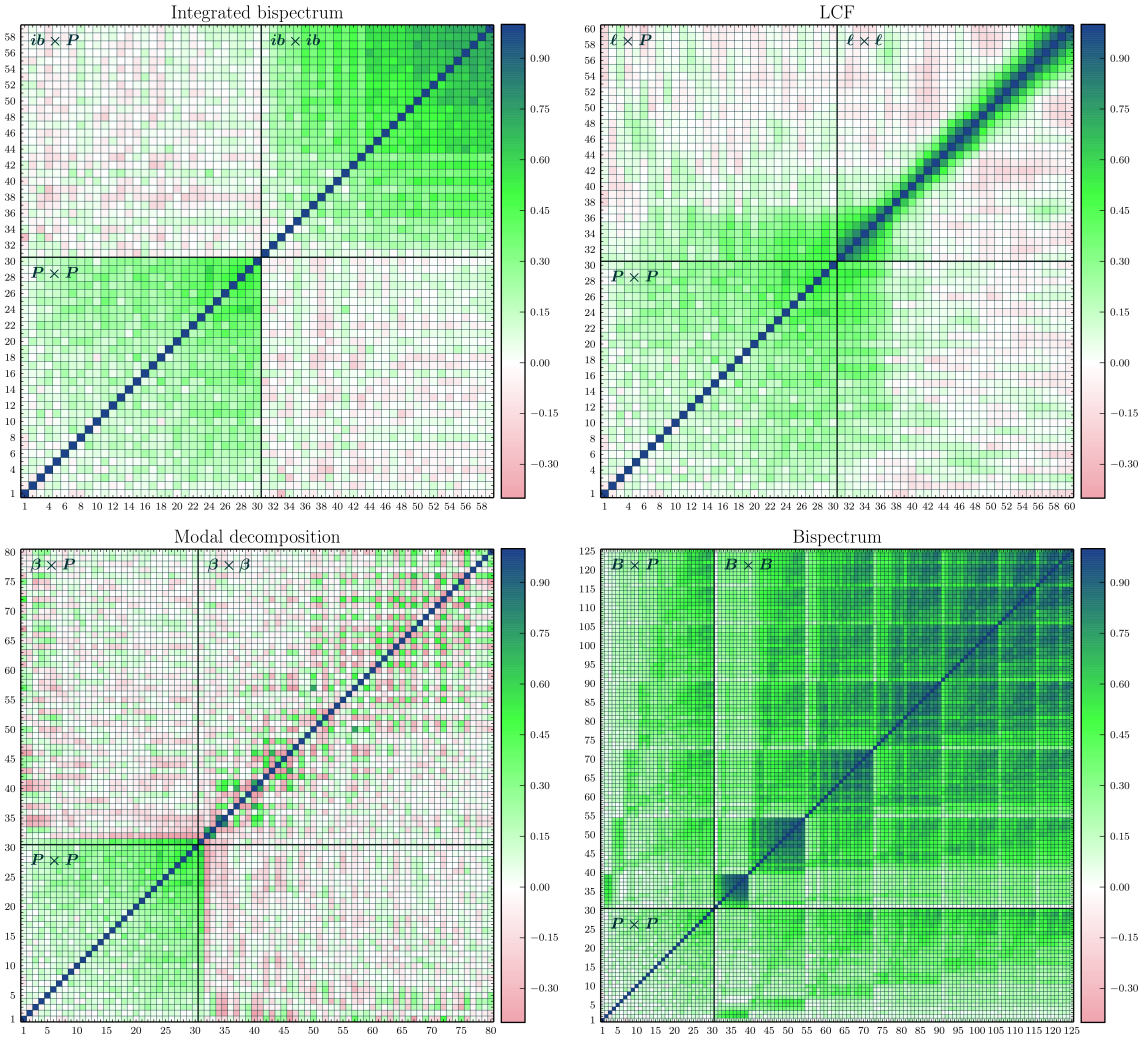


Fig. 6.7 Correlation matrices for (clockwise from top left) $P + ib$, $P + \ell$, $P + B$ and $P + \beta$ at redshift $z = 0.0$. In each panel, the lower-left quadrant contains the power spectrum auto-correlation ($P \times P$), while the upper-right quadrant contains the auto-correlation of the corresponding 3-point correlation measure. The upper-left and lower-right quadrants contain the cross-covariance.

In Sections 6.6 and 6.7 we quantify the impact of these complex non-diagonal covariances on estimates of signal-to-noise and Fisher forecasts.

Correlation matrices

We plot correlation matrices for the measurements $P + ib$, $P + \ell$, $P + \beta$ and $P + B$ in Fig. 6.7. We show measurements only at $z = 0$ where differences between the Gaussian and non-Gaussian covariances are largest.

The correlation coefficient $\mathbf{r}_{ij}$ between two data bins i and j is defined to satisfy $\mathbf{r}_{ij} \equiv \hat{\mathbf{C}}_{ij}/\sqrt{\hat{\mathbf{C}}_{ii}\hat{\mathbf{C}}_{jj}}$, where $\hat{\mathbf{C}}$ is the covariance matrix estimated from the simulation suite,

$$\hat{\mathbf{C}}_{ij} = \frac{1}{N_{\rm real}} \sum_{n=1}^{N_{\rm real}} \left[\hat{S}_i^{(n)} - \hat{\bar{S}}_i\right]\left[\hat{S}_j^{(n)} - \hat{\bar{S}}_j\right], \tag{6.57}$$

and $N_{\rm real} = 200$ is the number of realisations. To measure an auto-covariance the data vector S contains all measurements of a single proxy, $S = (X_{a,1}, \ldots, X_{a,n})$ or to measure a cross-covariance it contains all measurements from a pair, $S = (X_{a,1}, \ldots, X_{a,n_1}, X_{b,1}, \ldots, X_{b,n_2})$, where $X_a, X_b \in \{P, B, \beta, ib, \ell\}$. The correlation matrix measures the degree of coupling between different measurements. Its elements take values between -1 (where the bins are fully anti-correlated) and $+1$ (where the bins are fully correlated). A value of zero corresponds to independent measurements. For comparison, the Gaussian covariance matrices for P, B, β and ib are diagonal, whereas for ℓ there are correlations between neighbouring bins with similar r because it is a real-space statistic and therefore includes contributions from many Fourier configurations. In the Gaussian approximation the cross-covariance between P and any bispectrum proxy is zero.

Integrated bispectrum and line correlation function

Correlation measurements for the integrated bispectrum appear in the top-left panel of Fig. 6.7. The $ib(k)$ measurements show stronger auto-correlations than $P(k)$ as k increases, while the $P \times ib$ cross-correlation is relatively featureless. This indicates that the two data sets are nearly independent. Similarly, we find that the $P \times \ell$ cross-correlation is nearly featureless except where the smallest r bins and highest k bins show significant correlation. Relative to the Gaussian covariance matrix for ℓ, the r bins with $r \lesssim 50\,h^{-1}$Mpc are more strongly correlated due to non-linear growth.

Modal bispectrum

In the lower-left panel of Fig. 6.7 we present measurements of the correlation coefficients for $P + \beta^R$. These have not previously been reported. As explained in Sec. 6.3.4 these measurements apply to the R-basis, for which the covariance matrix is *constructed* to be diagonal in the Gaussian approximation. We find that only the first two modes are correlated with the majority of $P(k)$ bins. This is reasonable because the lowest modes probe the most scale-independent features of the phase bispectrum. The remainder show low-to-moderate correlation or anti-correlation due to non-linear effects.

Fourier bispectrum

For $P+B$ (bottom-right panel of Fig. 6.7) the correlation matrix has an approximate block structure due to the ordering of the 95 triangle configurations that we measure. The blocks correspond to groups of adjacent configurations with shared values of k_1 or k_2. While the power spectrum $P(k)$ shows mild correlations between different bins at high k, the bispectrum exhibits much stronger correlations. There are also non-zero cross-correlations between power spectrum and bispectrum bins. The correlation between power spectrum and bispectrum tends to be higher when $P(k)$ and $B(k_1, k_2, k_3)$ have

wave number bins that overlap. Similarly, the correlation between different bispectrum bins is higher when the configurations share at least one wave number. However, even configurations that have no wave numbers in common can be strongly correlated, with correlation coefficient as large as ~ 0.8, due to non-linear growth.

Cross-covariances

Finally, we have computed the correlation matrices between the bispectrum and its proxies. These enable us to identify which bispectrum configurations contribute most to individual bins of ib, ℓ or β^R.

We find that B and ib are very weakly correlated, which we attribute to ib being dominated by more strongly squeezed triangles than any we include in the 95 measured configurations of B. The line correlation function is found to be correlated with a majority of bispectrum configurations when $r \lesssim 40\,h^{-1}\text{Mpc}$. This indicates that the line correlation function is sensitive to many different shapes of Fourier triangle. Finally, we find that the first two β^R modes are strongly correlated with the bispectrum over large range of triangles, while the remainder are generally more correlated with triangles on the largest scales (that is, lower triangle index). This structure is similar to the $P+\beta^R$ correlation matrix. We do not find particularly strong correlations for $\ell \times ib$, but $\ell \times \beta^R$ shows that the line correlation function at small r is highly correlated with the first two β^R modes. This is consistent with the observation that both are sensitive to a wide range of Fourier configurations.

6.6 Cumulative signal-to-noise of the bispectrum proxies

Before discussing the constraining power of each proxy we first compute the available signal-to-noise. This is an intermediate step that characterises the significance with which measurements of each proxy can be extracted from a data set. Negligible signal-to-noise would normally imply poor prospects for parameter constraints. For example, Chan and Blot (2017) and Kayo et al. (2013) studied the signal-to-noise as a proxy for the information content of the Fourier bispectrum in the context of large-scale structure and weak lensing, respectively.

6.6.1 Numerical procedure

The cumulative signal-to-noise $\mathscr{S}/\mathscr{N}$ up to a maximum wave number k_{max} is defined by

$$\left(\frac{\mathscr{S}}{\mathscr{N}}\right)^2 \equiv \sum_{k_i,k_j\leq k_{\text{max}}} S_i \mathbf{C}^{-1}_{ij} S_j, \tag{6.58}$$

where S is the vector of typical values for either a single proxy or a combination of proxies, defined below Eq. (6.57). In this and subsequent sections we drop the use of a hat to denote an estimated value, and an overbar to denote a mean. The sum in (6.58) runs over all bins containing wavenumbers that satisfy the condition $k \leq k_{\rm max}$. For the Fourier bispectrum a bin corresponds to a triplet of wavenumbers (k_1, k_2, k_3), all of which are required to be smaller than $k_{\rm max}$.

We use the non-Gaussian covariance matrix measured from simulations, described in Sec. 6.5.3, which we denote by C_*. Its inverse C_*^{-1} is not an unbiased estimator of C^{-1}. A simple prescription to approximately correct for this bias is to rescale C_*^{-1} by an Anderson–Hartlap factor (Anderson, 2003; Hartlap et al., 2007), which yields

$$\mathsf{C}^{-1} \approx \frac{N_{\rm real} - N_{\rm bin} - 2}{N_{\rm real} - 1} \mathsf{C}_*^{-1}, \tag{6.59}$$

where $N_{\rm real}$ is the number of realizations used to estimate the covariance matrix and $N_{\rm bin}$ is its dimensionality.[6] Care should be taken when computing the numerical inverse C_*^{-1}, especially for combinations of measurements with signals of widely disparate magnitude. To avoid issues associated with ill-conditioning we first compute the correlation matrix $\mathsf{r}_{*,ij} = \mathsf{C}_{*,ij} / \sqrt{\mathsf{C}_{*,ii}\mathsf{C}_{*,jj}}$, whose entries lie between -1 and $+1$. We determine the inverse r_{ij}^{-1} using a singular value decomposition and check that all singular values are above the noise. Finally, we compute the inverse covariance using

$$\mathsf{C}_{*,ij}^{-1} = \frac{\mathsf{r}_{*,ij}^{-1}}{\sqrt{\mathsf{C}_{*,ii}\mathsf{C}_{*,jj}}}. \tag{6.60}$$

6.6.2 Results

In Fig. 6.8 we plot the resulting signal-to-noise measurements for the integrated bispectrum, line correlation function, the quantity B_ϵ defined in (6.16)—and used in the construction of the line correlation function and the modal bispectrum—and the Fourier bispectrum. (The signal-to-noise from B_ϵ and the reconstructed modal bispectrum give almost identical results as, due to our choice of weighting function (see Eq. (6.24)), Eq. (6.8) implies $B_\epsilon(k_1, k_2, k_3) \approx \sum_n \beta_n^Q Q_n(k_1, k_2, k_3) / \sqrt{k_1 k_2 k_3}$.)

Each panel of Fig. 6.8 shows the cumulative signal-to-noise of the Fourier bispectrum or a proxy (blue circles), together with the power spectrum (black crosses) and their combination including the cross-covariance matrix (red stars). Our measurements of the integrated bispectrum and line correlation function carry forward the binning

[6] Although the Anderson–Hartlap prescription is simple to apply, it has been pointed out by Sellentin and Heavens (2016) that this rescaling simply broadens the Gaussian likelihood of the data. These authors argued that the distribution of the data is more accurately modelled by a t-distribution.

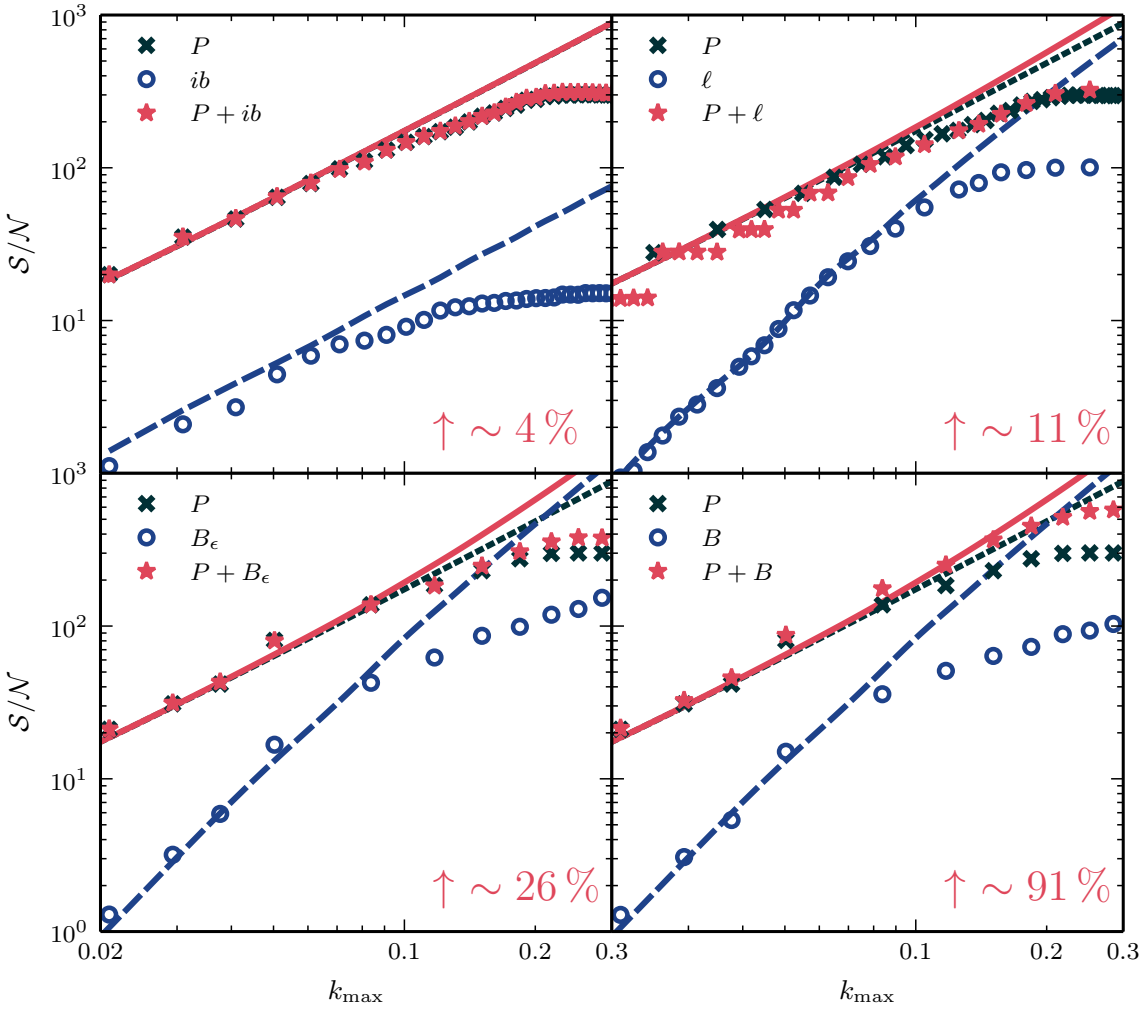


Fig. 6.8 Cumulative signal-to-noise at redshift $z = 0$ as a function of the maximal mode $k_{\max}$ for the measure X—equal to the integrated bispectrum, line correlation function, Fourier bispectrum or phase bispectrum (clockwise, starting from the top left panel). In each panel, blue circles refer to the measured signal-to-noise for X, while black crosses represent the signal-to-noise for the power spectrum. We plot the signal-to-noise for the combination $P + X$, including cross-covariance, as red stars. The blue, black and red lines give the theoretical prediction using the Gaussian approximation and tree-level SPT. The percentage quoted in the bottom right corner gives the increase in signal-to-noise relative to the power spectrum alone at $k_{\max} = 0.3\,h\,\mathrm{Mpc}^{-1}$.

procedure used in Sec. 6.5. The step-like structure that occurs for $P + \ell$ is due to a mismatch of scales between the power spectrum and the bins of the line correlation function. The first four data points in the B and B_ϵ panels use a bin size $\Delta k = 2k_\mathrm{f}$ in order to probe the low-k regime. The remainder derive from the measurements presented in Sec. 6.5 and use $\Delta k = 8k_\mathrm{f}$. In each panel, for comparative purposes, we plot lines of matching colour to show the signal-to-noise computed using a Gaussian approximation to the covariance matrix and tree-level SPT to evaluate any correlation measures it contains.

6.6.3 Discussion

First, we note that the Gaussian approximation overpredicts the signal-to-noise for each proxy X and its combination $P+X$ with the power spectrum. This is consistent with the results reported by Chan and Blot (2017). The over-prediction occurs because bins become coupled by non-linear evolution, and therefore do not provide independent information as the Gaussian approximation assumes. The effect can be quite severe: while the power spectrum signal-to-noise at $k_{\rm max} = 0.3\,h\,{\rm Mpc}^{-1}$ is over-predicted by a factor of three, the impact on the Fourier bispectrum and its proxies is much larger. In these cases the over-prediction ranges from a factor of ~ 5 or 8 for ib and ℓ up to more than an order of magnitude for the Fourier bispectrum. At smaller $k_{\rm max}$ the over-prediction is less, becoming significant for $k_{\rm max} \gtrsim 0.1\,h\,{\rm Mpc}^{-1}$.

The line correlation function, phase bispectrum and Fourier bispectrum *individually* contribute $\sim 30\%$ of the signal-to-noise of $P(k)$ at $k_{\rm max} = 0.3\,h\,{\rm Mpc}^{-1}$, while the integrated bispectrum achieves only 5% of the $P(k)$ signal-to-noise. For the Fourier bispectrum, this result is consistent with Chan and Blot (2017).

However, for estimating parameter constraints from the joint combination of P and B, or one of its proxies, the individual signal-to-noise contributed by one of these measurements is less important than whether it contains information that is not already present in the power spectrum. This is determined by the signal-to-noise of the combination $P+X$ compared to P alone. The different proxies show significant variation in the improvement from use of $P+X$, which we indicate as a percentage in the bottom-right corner of each panel. Although ℓ, B_ϵ and B individually carry roughly the same signal-to-noise, the uplift in $P+X$ varies from $\sim 11\%$ to $\sim 91\%$. Note that the signal-to-noise of $P+B$ receives a large improvement from the cross-covariance, which is in agreement with Chan and Blot (2017).

The discrepancy in uplift between B and B_ϵ is striking. If this discrepancy were to carry over to parameter constraints it would imply that the Fourier bispectrum carries *significantly* more constraining power than B_ϵ, even though both statistics are equivalent in the approximation of Gaussian covariance. If true, this would be very surprising. We return to this question in Sec. 6.7.5 after we have obtained forecast parameter uncertainties for B and its proxies, which enable us to precisely quantify the constraining power of each statistic.

6.7 Parameter uncertainty forecasts

In this section we collect our major results, which are Fisher forecasts of the error bars achievable on the parameter set $\theta_\alpha = (\Omega_m, \Omega_b, w_0, w_a, \sigma_8, n_s, h)$ of Table 6.1, based on a

fiducial flat ΛCDM cosmology. We perform these forecasts with and without inclusion of the bias parameters (b_1, b_2).

In Sec. 6.7.1 we summarise our implementation of the Fisher forecasting method, and in Sec. 6.7.2 we present and compare the forecasts from each proxy. By comparing forecasts with and without non-Gaussian covariances, and using different theoretical models to describe the dark matter density, we are able to characterise their influence on the final parameter constraints. These discussions appear in Sections 6.7.3 and 6.7.4, respectively. Finally, we return to the discussion of Sec. 6.6 and examine to what extent the signal-to-noise provides a reliable metric by which to estimate improvements in parameter constraints (Sec. 6.7.5).

6.7.1 Forecasting method

The Fisher formalism can be used to forecast the precision with which cosmological parameters could be measured in a future survey. Consider a data vector $\boldsymbol{x}$ containing measurements of any combination of statistical quantities. The likelihood function $\mathscr{L}(\boldsymbol{\theta} \mid \boldsymbol{x})$ is defined to be the probability of the data given the parameters $\boldsymbol{\theta}$, so $\mathscr{L}(\boldsymbol{\theta} \mid \boldsymbol{x}) = P(\boldsymbol{x} \mid \boldsymbol{\theta})$. Then the Fisher matrix $\mathbf{F}_{\alpha\beta}$ satisfies

$$\mathbf{F}_{\alpha\beta} \equiv -\left\langle \frac{\partial^2 \ln \mathscr{L}(\boldsymbol{\theta} \mid \boldsymbol{x})}{\partial\theta_\alpha \partial\theta_\beta} \right\rangle . \tag{6.61}$$

The expected 1σ error on each parameter θ_α, marginalised over all other parameters, can be obtained from the diagonal elements of the inverse Fisher matrix using $\sigma^2(\theta_\alpha) = (\mathbf{F}^{-1})_{\alpha\alpha}$. To simplify the computation of $\mathbf{F}_{\alpha\beta}$ we make the assumption that the likelihood function is a multivariate Gaussian,

$$\mathscr{L} = \frac{1}{\sqrt{(2\pi)^n |\mathbf{C}|}} \exp\left[-\frac{1}{2} (\boldsymbol{x} - \boldsymbol{\mu})^\mathsf{T} \mathbf{C}^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) \right], \tag{6.62}$$

where T denotes a matrix transpose and $|\mathbf{C}| = \det \mathbf{C}$ is the determinant of $\mathbf{C}$. We have written the mean of the data vector as $\boldsymbol{\mu} = \langle \boldsymbol{x} \rangle$, and its covariance matrix is $\mathbf{C}_{ij} = \langle x_i\, x_j \rangle - \mu_i\, \mu_j$. Testing the validity of this assumption to high precision would require many more independent simulations, but estimates of the bispectrum PDF based on our 200 boxes have not shown any significant deviations from Gaussianity, and we obtained a similar result for each of its proxies (see Sec. 5.6.1 for a discussion on the PDF of the LCF). Moreover, as shown in Hahn et al. (2018), even a significant non-Gaussian likelihood does not necessarily lead to a strong impact on the derived parameter constraints. Thus, continuing with a Gaussian likelihood, it can be shown that (Tegmark, 1997),

$$\mathbf{F}_{\alpha\beta} = \frac{1}{2} \mathrm{tr}\left[\mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial\theta_\alpha} \mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial\theta_\beta} \right] + \frac{\partial \boldsymbol{\mu}^\mathsf{T}}{\partial\theta_\alpha} \mathbf{C}^{-1} \frac{\partial \boldsymbol{\mu}}{\partial\theta_\beta}. \tag{6.63}$$

The first term measures variation of the covariance matrix with respect to the parameters, which is often a smaller effect than the variation of the means represented by the second term. In the approximation that this first term may be neglected the Fisher matrix can be computed in terms of the inverse covariance matrix for the fiducial model. Our procedure to obtain this matrix from the simulation suite has already been described in Sections 6.5 and 6.6.

Survey configuration

The Fisher formalism depends explicitly on details of the survey under discussion, both through the specification of the data vector $\boldsymbol{x}$—such as how many redshift bins are used and which Fourier configurations are included—and the properties of the covariance matrix $\mathbf{C}$. In the following we adopt the parameters of an idealised survey of large-scale structure consisting of three independent redshift slices at $z = 0$, $z = 0.52$ and $z = 1$. Each slice has volume $V = 3.375\,h^{-3}\,\mathrm{Gpc}^3$ and a mode cutoff at $k_{\mathrm{max}} = 0.3\,h\,\mathrm{Mpc}^{-1}$. The total Fisher matrix can be written as a sum of the Fisher matrix in each slice,

$$\mathbf{F}^{\mathrm{LSS}}_{\alpha\beta} = \mathbf{F}_{\alpha\beta}(z=0) + \mathbf{F}_{\alpha\beta}(z=0.52) + \mathbf{F}_{\alpha\beta}(z=1). \tag{6.64}$$

We assume that, in each redshift bin, the number density of galaxies is sufficiently high that the effect of shot noise is small. We do not include redshift-space distortions, the effect of complex survey geometry or the influence of super-survey modes. In general, all of these effects will be significant for a realistic survey and cannot be neglected. However, in this chapter our intention is to address the question of whether the proxies described in Sec. 6.2 can be competitive with measurements of the Fourier bispectrum *in principle*. Survey-specific effects will generally reduce the number of configurations that can be measured, or increase the noise on those for which measurements are possible. This will typically weaken the performance of the proxies, meaning that their neglect gives us an estimate of the best-case scenario. While we do not anticipate that astrophysical or observational systematics will affect any one proxy more than the others, this is an interesting question to explore in future.

Each of the constraints we present includes a prior from the cosmic microwave background power spectrum. We implement this prior by adding a fourth Fisher matrix,

$$\mathbf{F}^{\mathrm{tot}} = \mathbf{F}^{\mathrm{LSS}} + \mathbf{F}^{\mathrm{CMB}}. \tag{6.65}$$

Details of the computation of $\mathbf{F}^{\mathrm{CMB}}$ for our choice of fiducial parameters were given by Smith et al. (2014).

6.7.2 Constraining power of the bispectrum and its proxies

In this section we present our forecasts. To minimise modelling errors we construct the Fisher matrix for each proxy using quantities measured from simulation, except for derivatives with respect to the bias parameters which cannot be obtained in this way. For the Fourier bispectrum we compute these derivatives analytically by differentiating the one-loop power spectrum (6.36) and the tree-level bispectrum (6.37). Once the derivatives have been obtained we replace occurrences of the dark matter power spectrum and bispectrum with their measured values. Our prescription for the proxies is similar, using the one-loop power spectrum to estimate derivatives of $P(k)$ and tree-level formulae together with the formulae of Sec. 6.3 to estimate derivatives of the proxy.

We plot the forecast 1σ confidence contours in Fig. 6.9. Each panel shows predicted joint constraints for a pair of parameters after marginalising over all the others. The grey shaded region marks the constraint predicted from measurements of the power spectrum only, except for inclusion of the CMB prior that we apply to all estimates. The solid dark-blue line marks the constraint predicted from $P + ib$; the long-dashed red line marks the constraint predicted from $P + \ell$; the short-dashed light-blue line marks the constraint predicted from $P + \beta$; and the solid black line marks the constraint predicted from $P + B$. We summarise the marginalised 1σ error bars in Table 6.3. The value in parentheses following each uncertainty indicates the percentage improvement compared to use of $P(k)$ alone.

Improvement from three-point correlation data

First consider the joint constraints from $P + B$ (solid black lines in Fig. 6.9). These demonstrate that substantial improvements can be achieved compared to measurement of the power spectrum only. This is especially evident for σ_8 and the two bias parameters, for which the improvement is roughly 70%–80%; compare this with the second column of Table 6.3. This is perhaps unsurprising: the bispectrum constrains a different combination of σ_8 and b_1 than the power spectrum, and therefore assists in breaking their degeneracy (Fry, 1994; Matarrese et al., 1997). Nevertheless, other parameters that do not participate in this degeneracy also experience improvements in the range 13%–22%, with the exception of Ω_b. This is already very well-measured by the CMB prior, and large-scale structure measurements can add little new information. These conclusions are similar to those reported by Sefusatti et al. (2006), who suggested that inclusion of Fourier bispectrum measurements could reduce uncertainties on Ω_m and σ_8 by a factor in the range 1.5 to 2.

Next, the forecast for the integrated bispectrum (solid dark-blue lines) shows that it offers negligible improvement, of order $\sim 2\%$, in comparison to P alone. This is

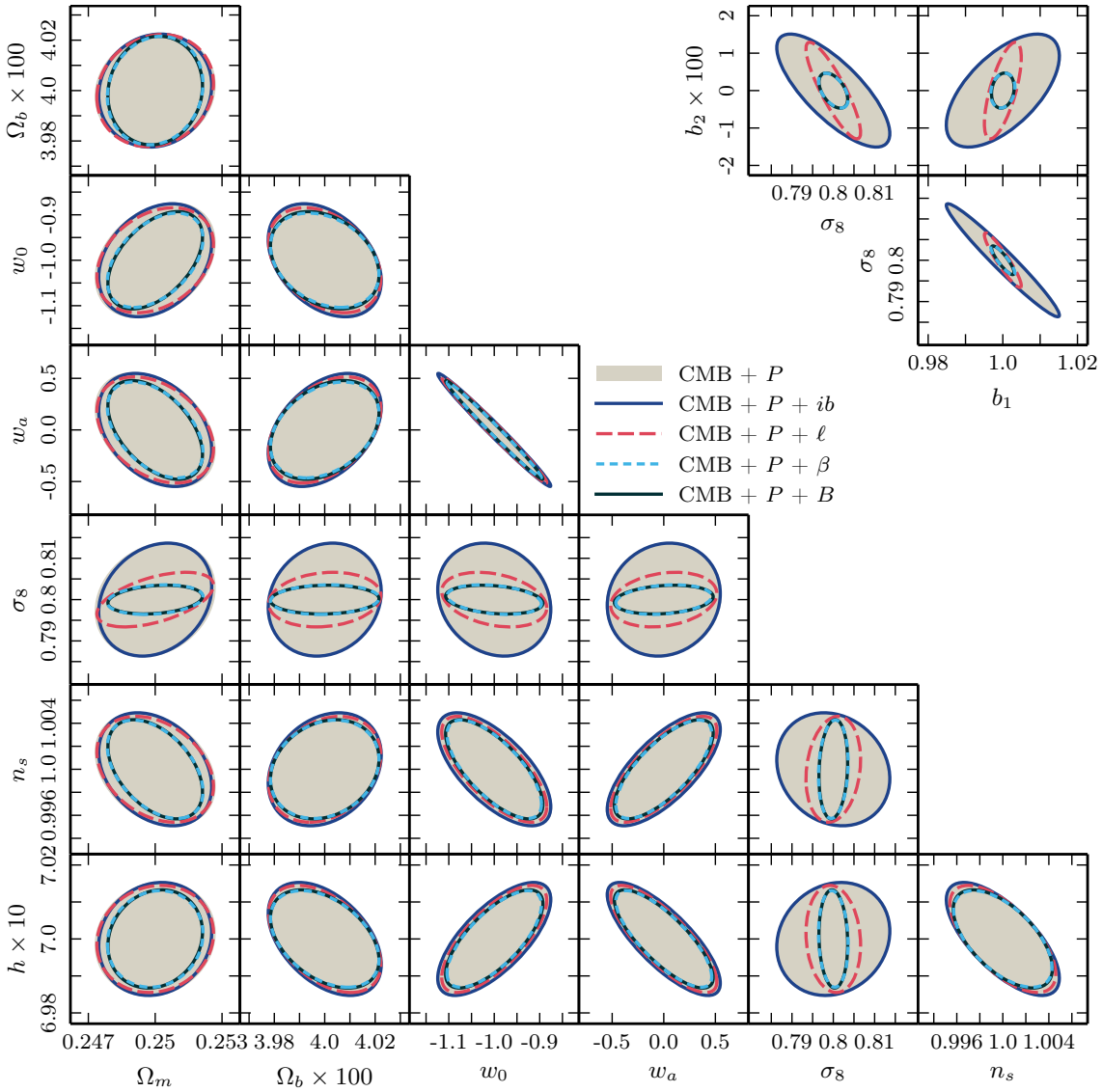


Fig. 6.9 Comparison of marginalised 1σ likelihood contours forecast from a combination of the power spectrum and one of the following 3-point measures: integrated bispectrum (dark blue, solid), line correlation function (red, long-dashed), modal decomposition (light blue, short-dashed) and Fourier bispectrum (black, solid). The grey shaded regions show the error ellipses for the power spectrum alone. All forecasts include priors from a Planck-like CMB experiment and use a cut-off scale $k_{\rm max} = 0.3\,h\,{\rm Mpc}^{-1}$. The covariance matrices and parameter derivatives for the Fisher forecasts shown here are all derived from our simulation results in Sec. 6.5.

consistent with the very small dependence on cosmological parameters discussed in Sec. 6.5.2, and the low signal-to-noise obtained in Sec. 6.6. On the other hand, the line correlation function offers comparable constraints to the Fourier bispectrum for σ_8 and b_1, which receive improvements of 53% and 68%, respectively. Eggemeier and Smith (2017) demonstrated that this occurs because the line correlation function is nearly independent of b_1 and therefore probes a different direction in parameter space than P or B. Also, inclusion of ℓ measurements increases sensitivity to the dark energy

parameters w_0 and w_a by $\sim 9\%$. These improvements are only marginally degraded compared to those from $P+B$, which are of order 15%.

Finally, Fig. 6.9 demonstrates that the modal bispectrum with $n_{\max} = 50$ (short-dashed light-blue lines) is predicted to yield error bars nearly equivalent to the Fourier bispectrum with 95 triangles. Indeed the modal estimator for some parameters does slightly better than the bispectrum, which may lose some information due to the large bin width $\Delta k = 8k_{\rm f}$. Nevertheless, there is no sign of the significant difference in constraining power between B and B_ϵ—which is the quantity implicitly measured by β with our choice of basis—that was suggested by our analysis of signal-to-noise in Sec. 6.6. We return to this apparent discrepancy in Sec. 6.7.5 below. Just as important, the differences between the cases $n_{\max} = 10$ and $n_{\max} = 50$ are mostly negligible. Therefore, even with as few as $n_{\max} = 10$ modes, the modal decomposition retains nearly the full constraining power of the bispectrum. However, it should be remembered that Fig. 6.5 suggests the Fourier bispectrum reconstructed with so few modes will introduce more significant scatter. In a realistic analysis, these reconstruction errors could manifest themselves as a bias on the best-fit cosmological parameters. Unfortunately we cannot account for this bias in our Fisher analysis, but it deserves further investigation.

Combination with other observables

The strong degeneracy between σ_8 and b_1 can be broken by other means. For example, it is possible to use weak lensing measurements that probe the matter power spectrum directly. Given that inclusion of 3-point correlation data yields the largest improvements for σ_8 and the bias, it is worthwhile considering what improvements should be expected were the bias to be fixed by other cosmological observations.

In a scenario of this kind the power spectrum constraints would not be weakened by marginalisation over the bias parameters, and therefore inclusion of 3-point correlation data would no longer yield such a dramatic improvement for σ_8. However, we still find encouraging improvements for many parameters. For example, inclusion of either Fourier or modal bispectrum measurements would decrease uncertainty on σ_8 by $\sim 25\%$ and all other parameters except Ω_b by 10%–15%. Inclusion of ℓ measurements would decrease uncertainty on σ_8 by 20%, on the dark energy parameters by $\sim 10\%$, and for all other parameters by $\lesssim 5\%$. We conclude that, even in the extreme case that b_1 and b_2 can somehow be determined exactly, inclusion of 3-point correlation data still provides valuable additional information.

These Fisher forecasts should be interpreted with some care. As explained above, we do not include a number of astrophysical and observational effects that complicate the analysis of realistic galaxy survey data. These include redshift uncertainties, redshift-space distortions, irregular survey geometries and shot noise. In particular, for the forecasts presented here the effective shot noise is set by the number density $\bar{n} =$

Table 6.3 Marginalised 1σ parameter uncertainties for the power spectrum and its combination with a 3-point correlation measure, including CMB priors. All quoted values are derived from the measured covariance matrices and parameter derivatives with $k_{\max} = 0.3\,h\,\mathrm{Mpc}^{-1}$. The percentages in parentheses refer to the improvement over the P-only results.

	P	$P+B$		$P+\beta$ $n_{\max}=$ 50		$P+\beta$ $n_{\max}=$ 10		$P+\ell$		$P+ib$	
Ω_m	0.00179	0.00140	(22 %)	0.00141	(21 %)	0.00144	(19 %)	0.00172	(4 %)	0.00167	(7 %)
Ω_b	0.00015	0.00014	(5 %)	0.00014	(5 %)	0.00014	(4 %)	0.00015	(2 %)	0.00015	(1 %)
w_0	0.084	0.070	(16 %)	0.068	(19 %)	0.069	(17 %)	0.076	(9 %)	0.082	(2 %)
w_a	0.370	0.315	(15 %)	0.306	(17 %)	0.310	(16 %)	0.338	(9 %)	0.360	(3 %)
σ_8	0.0092	0.0023	(75 %)	0.0024	(74 %)	0.0025	(73 %)	0.0043	(53 %)	0.0090	(2 %)
n_s	0.00327	0.00284	(13 %)	0.00281	(14 %)	0.00284	(13 %)	0.00303	(7 %)	0.00323	(1 %)
h	0.00103	0.00087	(15 %)	0.00086	(16 %)	0.00087	(15 %)	0.00095	(7 %)	0.00101	(2 %)
b_1	0.0103	0.0020	(81 %)	0.0021	(79 %)	0.0022	(79 %)	0.0032	(68 %)	0.0100	(3 %)
b_2	0.0100	0.0031	(69 %)	0.0031	(69 %)	0.0031	(69 %)	0.0085	(15 %)	0.0100	(1 %)

$0.125\,h^3\,\mathrm{Mpc}^{-3}$ of particles in our simulation suite. This is substantially larger than the galaxy number densities that will be achieved by upcoming surveys. We return to this issue in Sec. 6.8.2, where we discuss how our predictions would be modified by a more realistic number density.

6.7.3 Effect of non-Gaussian covariance and cross-covariance

The non-Gaussian covariance measured in simulations differs from the Gaussian approximation in two ways: (1) it includes additional contributions to the variance of each bin from higher-order correlations, and (2) it adds or enhances coupling between different bins of a single proxy, and between bins of different proxies. These non-Gaussian corrections generally lead to weaker parameter constraints when compared to forecasts constructed using the Gaussian approximation, because this assumes that every bin contributes independent information. In this section we compare the relative impact of non-Gaussian covariance for the different proxies by contrasting Fisher forecasts made with and without its inclusion. We give results for the combinations $P + ib$, $P + \ell$, $P + \beta$ and $P + B$ and each choice of theoretical model—tree-level SPT, 1-loop SPT, or the halo model. For all forecasts in this work with (or without) non-Gaussian covariances, the non-Gaussian contributions to the covariance are always included (or ignored) in both P and the higher-order correlation statistic at the same time.

Increase in uncertainty from non-Gaussian contributions

Fig. 6.10 shows the relative increase $\sigma_{NG}/\sigma_G - 1$ in predicted uncertainty for each parameter when non-Gaussian contributions are included. To estimate σ_G we use the expressions for Gaussian covariance given in Sec. 6.3 with each quantity replaced by its value measured from our simulations. For example, to construct the Gaussian covariance for ib we use Eq. (6.14) with σ_L^2 replaced by its measured value. We could equally well have constructed similar estimates using one of the theoretical models to calculate such values, but the result is not very different. The discussion in this section would continue to apply if we were to reproduce Fig. 6.10 using estimates generated by any of these prescriptions.

The increase in uncertainty induced by inclusion of non-Gaussian effects depends on the measure of 3-point correlations used to generate constraints, the method used to estimate the Gaussian covariance matrix, and the parameter in question. In general we find that the Gaussian approximation under-predicts the uncertainty for the Fourier bispectrum more strongly than for its proxies. Note also that—although $P + \beta$ and $P + B$ yield nearly identical constraints when the non-Gaussian covariance is used, as described in Sec. 6.7.2—the importance of the non-Gaussian covariance for these combinations is not the same. Since the quantity B_ϵ measured by β is not the same as

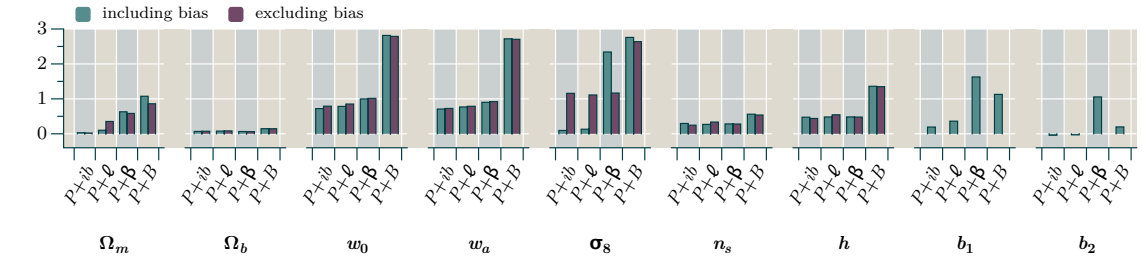


Fig. 6.10 Increase in parameter uncertainties from non-Gaussian covariances, measured using $\sigma_{NG}/\sigma_G - 1$, where σ_{NG} (σ_G) is the predicted error bar using the non-Gaussian (Gaussian-like) covariance from simulations. Predictions that include (do not include) a marginalisation over the bias parameters are in blue-green (purple).

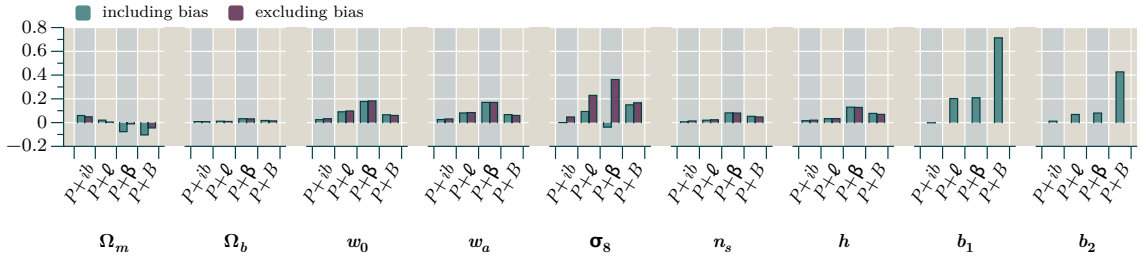


Fig. 6.11 Improvement in parameter uncertainties from the inclusion of cross-covariance, measured using $\sigma_{NG-\text{no}-CC}/\sigma_{NG-\text{with}-CC} - 1$, where $\sigma_{NG-\text{with}-CC}$ is the error bar predicted using non-Gaussian covariance measured from simulations and $\sigma_{NG-\text{no}-CC}$ is the error bar predicted from the same covariance matrix, except with cross-covariances between P and each 3-point statistic set to zero. Predictions that include (do not include) a marginalisation over the bias parameters are in blue-green (purple).

B, neglecting cross-covariance with P (as the Gaussian covariance does) will leave out different information for $P+\beta$ compared to $P+B$.

Inclusion of non-Gaussian covariance impacts uncertainties for w_0, w_a and σ_8 more significantly than the other parameters. This non-uniformity means that it is not obvious how inclusion of non-Gaussian covariance might impact constraints from 3-point correlations on further parameters not considered here. For instance, a number of authors have used Gaussian covariances to forecast future constraints on a primordial bispectrum generated by inflation; see Scoccimarro et al. (2004), Sefusatti and Komatsu (2007), Sefusatti et al. (2012), Baldauf et al. (2016), Welling et al. (2016) and Tellarini et al. (2016). It is not yet clear how these forecasts will change when more realistic non-Gaussian covariances are used.

Inclusion of cross-covariance

In Fig. 6.11 we summarise the influence of cross-covariance between P and the 3-point measures by comparing constraints using the full non-Gaussian covariance to

constraints where the cross-covariance has been set to zero. We find that inclusion of cross-covariances *reduces* the predicted uncertainties for nearly all parameters and choices of combination $P + X$, whether or not we marginalise over galaxy bias. In the few cases where inclusion of cross-covariance did not reduce the uncertainties (e.g. constraints on Ω_m from $P + B$ and $P + \beta$), the predicted error bar is weakened by less than 12% of the error bar without cross-covariance. Overall, we find that ignoring cross-covariances can overestimate uncertainties by up to $\sim 40\%$ when we do not marginalise over the bias, and by $40 - 70\%$ for the special case of bispectrum constraints on the bias parameters themselves.

This reduction of uncertainties due to inclusion of cross-covariances may be surprising. While we have not explicitly identified the source of the improved constraining power, this is not a new feature of Fisher forecasts using non-Gaussian covariances. For example, a number of authors using cross-correlations between cluster counts, weak lensing power spectra and the weak lensing bispectrum have found that parameter constraints can improve when cross-covariances between strongly-coupled measurements are included (Kayo et al., 2013; Sato and Nishimichi, 2013; Takada and Bridle, 2007). But it is also possible that our improvements are partly due to the galaxy biasing model we have chosen. A simulation of halos, rather than dark matter alone, could be used to verify the effect when simultaneously constraining both cosmological parameters and galaxy bias.

The conclusion of this discussion is that an accurate estimate for the covariance matrix, including non-Gaussian contributions and off-diagonal terms, is important if we wish to obtain reliable constraints. Unfortunately, this is especially true for the Fourier bispectrum for which the Gaussian approximation most significantly underestimates the true parameter uncertainties. This implies that surveys aiming to generate constraints from inclusion of B measurements cannot evade the computational difficulties associated with estimating their covariance matrix.

To mitigate these difficulties we could consider use of $P + \beta$ rather than $P + B$. As we have seen in Sec. 6.7.2, these combinations yield nearly equivalent constraints using 95 Fourier configurations and 50 modal coefficients respectively, and therefore the modal decomposition makes the information content of the bispectrum more accessible by reducing the size of the covariance matrix needed to obtain it. We consider the efficiency with which each proxy can compress the information carried by B in Sec. 6.8.1.

6.7.4 Theory-dependence of the forecasts

In Sec. 6.7.2 we have presented our Fisher forecasts based on simulated data, and in Sec. 6.7.3 we have discussed the influence of non-Gaussian covariance and cross-covariances. These results enable us to assess the information content carried by the

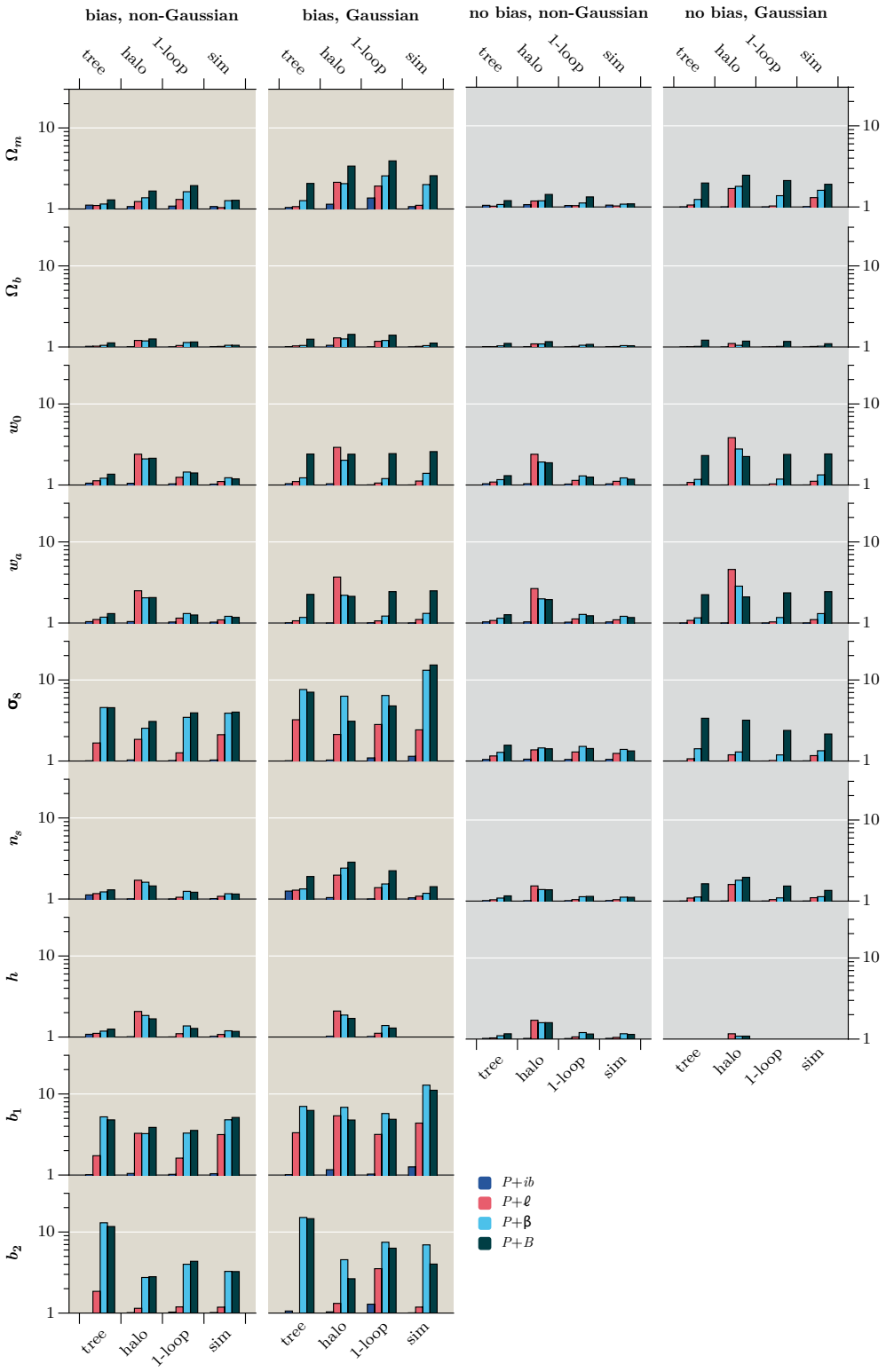


Fig. 6.12 Improvements in parameter uncertainties from the addition of 3-point statistics are shown as bars with height $\sigma(P)/\sigma(P+X)$, the ratio of parameter errors from P only and the combination $P+X$. The labels at the top of each column indicate whether bias parameters are included or excluded, and whether Gaussian or non-Gaussian covariances are used. Each group of four bars corresponds to a different choice of theoretical model, and the colour of each bar indicates the $P+X$ combination. We note that, since the tree-level power spectrum does not depend on b_2, for the two tree-level bar groups in the last row, the bar heights measure $\sigma_{\rm max}/\sigma(P+X)$, where $\sigma_{\rm max}$ is the maximum error on b_2 among the four $\sigma(P+X)$ values.

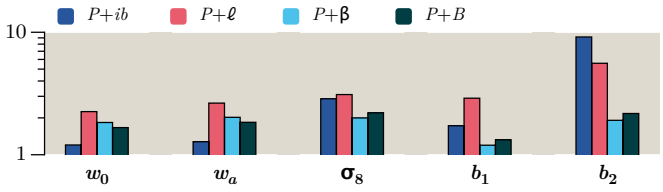


Fig. 6.13 Sensitivity factors, defined as the ratio between the largest and smallest forecast parameter uncertainty among the three theoretical models, for each $P + X$ combination. The forecasts compared here include bias parameters and use non-Gaussian covariances.

Fourier bispectrum and its proxies, but the question of how easily these statistics can be deployed remains open. In particular, we would like to know whether the use of simulated data is essential, or whether any of the models described in Sec. 6.3 are sufficient. In this section we study the dependence of our forecasts on the choice of theoretical model used to estimate the derivatives $\partial\boldsymbol{\mu}/\partial\theta_\alpha$ in Eq. (6.63).

Match to forecast from simulations

First, we consider whether there is a model that provides a clear best-match to the forecast using simulated data. Fig. 6.12 compares the forecasts for each parameter using different prescriptions for the covariance matrix and for different choices of theoretical model, with marginalisation over the bias included or excluded. The bar heights represent the reduction in the predicted uncertainty provided by a given combination, relative to the base model of power spectrum data only combined with a CMB prior. The results of Sec. 6.7.2 are labelled 'sim'. Unfortunately, for each combination $P + X$ there is no single choice of theoretical model yielding forecasts that provide the best match to the 'sim' outcome for all parameters—with or without marginalisation over bias.

For example, consider the combination $P + B$ in the first column of Fig. 6.12. This summarises forecasts generated by including non-Gaussian covariance and marginalisation over the bias. For σ_8 it is 1-loop SPT that gives the best match to the 'sim' result, but for the linear bias parameter b_1 the best match comes from tree-level SPT.

Alternatively, one could ask whether any one model provides uniformly conservative or uniformly optimistic forecasts. If so, that model could be used to estimate upper or lower limits on the uncertainty for any chosen parameter. But Fig. 6.12 demonstrates that there are no models with such properties. For example, focusing again on the first column, there is no single choice of theoretical model for $P + B$ that forecasts the largest or smallest improvement for all parameters.

Sensitivity to theory error

Next, we study the variation in forecasts for the Fourier bispectrum and its proxies when we change the model used to compute $\partial\boldsymbol{\mu}/\partial\theta_\alpha$. To understand the sophistication required to obtain accurate models we will need to understand which of these statistics (if any) are especially sensitive or immune to theoretical mis-modelling. We measure this dependence by a *sensitivity factor*, which we define to be the ratio between the largest and smallest forecast uncertainties taken over the models of Sec. 6.3. A sensitivity factor close to unity indicates that a forecast uncertainty depends only weakly on the choice of theoretical model, while a large value indicates that the model has a strong influence on the final outcome.

We plot these sensitivity factors in Fig. 6.13, computed with inclusion of all bias parameters and using non-Gaussian covariances. Therefore the sensitivity factor solely reflects the variation in uncertainty produced by different choices for theoretical model. We conclude that there is no single measure of 3-point correlations that consistently yields the largest or smallest sensitivity to variations in modelling. Therefore, there is apparently no single combination $P+X$ that should be preferred to minimise the effect of theory errors on inferred parameter constraints.

Ranking by constraining power

Neither of these criteria provide a rationale to prefer a choice of theoretical model. Nevertheless, we do find some general trends. Irrespective of theoretical model, we find the largest reductions in parameter uncertainties when the bias is constrained simultaneously with the cosmological parameters. Also, the Fourier bispectrum and modal bispectrum consistently offer the most significant improvements compared to P-only measurements, with very similar predicted uncertainties. The line correlation function achieves moderate improvement compared to P-only, while the integrated bispectrum has very weak constraining power—at least for the parameter set we consider. We conclude that $P+B$ or $P+\beta$ should be preferred for constraints on ΛCDM parameters, with $P+\beta$ offering similar information at reduced computation cost as discussed at the end of Sec. 6.7.3.

Relative importance of modelling and non-Gaussian covariance

Finally, we consider the relative importance of non-Gaussian covariance and theoretical modelling for obtaining quantitatively accurate forecasts. In Fig. 6.14 we show the fractional difference in Fisher forecasts induced by variation of theoretical model (orange bars) and use of the Gaussian approximation (blue bars). To quantify the significance of theoretical modelling we plot $\max(|\sigma_{NG,i}/\sigma_{NG}(\text{sim})-1|)$, where $i\in\{\text{tree},\text{1-loop},\text{halo}\}$. Therefore larger orange bars reflect more significant deviation from the simulated fore-

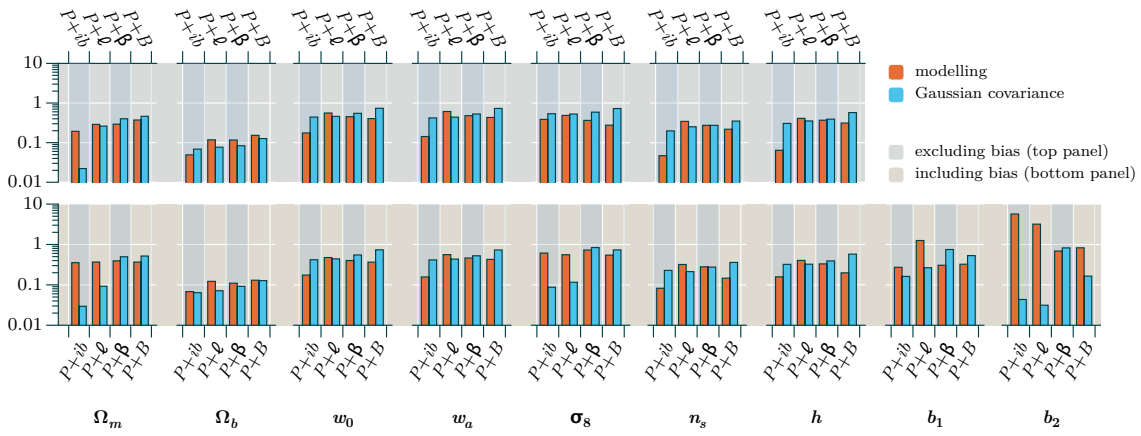


Fig. 6.14 Fractional difference in predicted uncertainties induced by theoretical modelling of derivatives (orange) or by using a Gaussian approximation to the covariance (blue). (See text for details of how the fractional differences are defined.)

cast due to theoretical uncertainty. Meanwhile we quantify the role of the covariance matrix by plotting $|\sigma_G(\text{sim})/\sigma_{NG}(\text{sim}) - 1|$, so increasing blue bars show that the Gaussian approximation generates more significant errors in the forecast.

Fig. 6.14 shows that the impact of theoretical uncertainty for $P+\beta$ and $P+B$ is generally less significant than neglect of non-Gaussian covariance, whether or not we marginalise over the bias. In contrast, for $P+\ell$ the effect of modelling nearly always dominates because of the difficulties with the halo model discussed in Sec. 6.5.2. For $P+ib$ the non-Gaussian covariance plays an important role if the bias parameters are not included, but theoretical modelling dominates when they are.

On balance, these results indicate that our forecasts are slightly less sensitive to theory error than to the approximation of Gaussian covariance. This could be because the inverse covariance weighting suppresses contributions from the non-linear regime where the theoretical predictions are most discrepant. But the difference is not large: the average variation in our predicted uncertainties from $P+B$ and $P+\beta$ due to theory modelling is 36%, whereas the variation due to Gaussian covariances is 49%. Therefore, we conclude that both issues must be addressed in order to obtain quantitatively accurate results. Finally, we note that the Fisher forecast method does not allow us to estimate the amount of bias on the recovery of a particular parameter stemming from an inaccurate theory model. This effect can be as problematic as an over- or underestimation of the parameter uncertainties, and must be carefully tested for each model using mock data.

6.7.5 Signal-to-noise as a proxy for the information content

It is now necessary to address the question of why the large discrepancy in uplift between the signal-to-noise of B and B_ϵ (equivalently β) observed in Sec. 6.6 did not translate into significant differences in the forecast for parameter uncertainties in Sec. 6.7.2.

Consider a vector of values S combining measures P and X of the 2- and 3-point correlation data, respectively, as defined below Eq. (6.57). For a given parameter θ the reduction in uncertainty compared to measurements from P alone can be estimated in the Fisher framework by

$$\frac{\mathbf{F}_\theta(S)}{\mathbf{F}_\theta(P)} = \sum_{i,j} \frac{\partial S_i}{\partial\theta} \mathbf{C}^{-1}_{ij} \frac{\partial S_j}{\partial\theta} \Big/ \sum_{i,j} \frac{\partial P_i}{\partial\theta} (\mathbf{C}^{\rm P})^{-1}_{ij} \frac{\partial P_j}{\partial\theta}. \tag{6.66}$$

To avoid ambiguity we use the notation $\mathbf{C}^{\rm P}$ to denote the covariance matrix of the power spectrum *only*. Meanwhile, the increase in signal-to-noise in the same scenario is given by

$$\frac{(\mathscr{S}/\mathscr{N})^2_S}{(\mathscr{S}/\mathscr{N})^2_P} = \sum_{i,j} S_i \mathbf{C}^{-1}_{ij} S_j \Big/ \sum_{i,j} P_i (\mathbf{C}^{\rm P})^{-1}_{ij} P_j. \tag{6.67}$$

The uplift in signal-to-noise is often taken as an approximation to the reduction in parameter uncertainty, which avoids the need to compute $\partial S_i/\partial\theta$. As we have seen in Sec. 6.5.2, these derivatives can be rather fragile and are susceptible to significant errors caused by theory mis-modelling. Unfortunately, when applied to $S = P + B$ and $S = P + \beta$ our analysis demonstrates that the ratios $\mathbf{F}_\theta(P+B)/\mathbf{F}_\theta(P)$ and $\mathbf{F}_\theta(P+B_\epsilon)/\mathbf{F}_\theta(P)$ are nearly equal, whereas the same ratios constructed using $\mathscr{S}/\mathscr{N}$ are very discrepant. Therefore we must conclude that improvements in signal-to-noise cannot always be interpreted as a predictor of the improvement in Fisher information.

Invariance of the Fisher matrix

First consider the Fisher matrix. Suppose we perform a redefinition so that $S_i \to S'_i = S'_i(S_j)$, where S'_i may be an arbitrary nonlinear function of the original measurements. For example, the transformation from B to B_ϵ is of this type. The derivative $\partial S_i/\partial\theta_\alpha$ transforms 'contravariantly' on its index i, in the sense $\partial S'_i/\partial\theta_\alpha = \sum_m (\partial S'_i/\partial S_m)(\partial S_m/\partial\theta_\alpha)$. Meanwhile, the covariance matrix becomes

$$\mathbf{C}^S_{ij} \to \mathbf{C}^{S'}_{ij} = \langle (S'_i - \bar{S}'_i)(S'_j - \bar{S}'_j)\rangle = \sum_{m,n} \frac{\partial S'_i}{\partial S_m} \frac{\partial S'_j}{\partial S_n} \mathbf{C}^S_{mn} + \cdots, \tag{6.68}$$

where '$\cdots$' denotes terms involving higher order correlations that we have not written explicitly. Provided these are small compared to the $\mathbf{C}^S_{mn}$ term, Eq. (6.68) shows that the covariance matrix also transforms 'contravariantly', and therefore that its inverse transforms 'covariantly'. Subject to these approximations we conclude that the Fisher matrix should be roughly *invariant*. This agrees with our observation that $\mathbf{F}_\theta(P+B)$ and $\mathbf{F}_\theta(P+B_\epsilon)$ are nearly equal, demonstrated numerically in Table 6.3.

Now consider the signal-to-noise. Since S_i has neither a co- or contravariant transformation law, the combination $\sum_{i,j} S_i \mathbf{C}^{-1}_{ij} S_j$ appearing in the signal-to-noise will typically

Table 6.4 Percent improvement of unmarginalised constraints using $P+B$ compared to P only at $z=0$.

Ω_M	Ω_B	w_0	w_a	σ_8	n_s	h	b_1	b_2
12.9 %	19.4 %	26.0 %	27.0 %	26.4 %	15.1 %	15.6 %	42.4 %	43.4 %

not be invariant. Therefore different choices S_i and S'_i may yield inequivalent results for $\mathscr{S}/\mathscr{N}$. For example, we have verified that using $P+\ln B$ predicts a significant increase in the signal-to-noise compared to $P+B$, whereas their Fisher matrices continue to agree. In Table 6.4 we summarise the improvement in unmarginalised constraints from the addition of B or B_ϵ. This demonstrates that empirically the increase in signal-to-noise from B_ϵ provides a more accurate estimate of the Fisher information than B. This property holds for both proxies of B_ϵ, namely the modal bispectrum, and the line correlation function.

Gaussian limit

This outcome is not inconsistent with the result that B and B_ϵ show an equivalent uplift in signal-to-noise in the Gaussian approximation. In this case the covariance matrix for B_ϵ is $\mathbf{C}^{B_\epsilon}_{ij} = \mathsf{N}\mathbf{1}_{ij}$, where the constant N takes the values 1, 2 or 6 for scalene, isosceles and equilateral configurations, respectively, as described in Sec. 6.3. In the same approximation the covariance matrix for the Fourier bispectrum is $\mathbf{C}^{B}_{ij} = \mathsf{N}P(k_{i_1})P(k_{i_2})P(k_{i_3})\mathbf{1}_{ij}$. Therefore we conclude that the signal-to-noise for B and B_ϵ is identically equal as

$$B_i(\mathbf{C}^B)^{-1}_{ij}B_j = B_{\epsilon i}(\mathbf{C}^{B_\epsilon})^{-1}_{ij}B_{\epsilon j} = \frac{1}{\mathsf{N}}\frac{B_i^2\mathbf{1}_{ij}}{P(k_{i_1})P(k_{i_2})P(k_{i_3})}. \tag{6.69}$$

In the Gaussian approximation the power spectrum is an independent source of information, which explains the agreement. However, once off-diagonal contributions in the covariance matrix are included, B and P are no longer independent and non-linear combinations may give very different results for the signal-to-noise.

6.8 Discussion

6.8.1 Compression and efficiency of the Fourier bispectrum proxies

In an ideal survey aiming to measure the Fourier bispectrum we should clearly choose a bin width Δk that is sufficiently small to reproduce all small-scale features of interest. However, because the number of Fourier configurations in a volume with mode cut-off $k_{\max}$ scales as $\sim (k_{\max}/\Delta k)^3$ this task will quickly become computationally expensive. And, as we have emphasised several times, a more serious problem is that we must

estimate and invert the covariance matrix for all these measurements. This requires us to perform at least as many N-body simulations as the number of configurations that we retain.

In this section we consider how well this large number of Fourier configurations can be compressed by the proxies described in Sec. 6.2. Suppose that available resources limit the number of simulations that can be performed in such a way that we can estimate an accurate covariance matrix for ~ 30 bins of the Fourier bispectrum or one of its proxies, in combination with another 30 measurements of the power spectrum $P(k)$. Among the measures of 3-point correlations that we consider, is there a preferred choice that provides optimal constraints on our set of cosmological parameters? If so, this measure would provide the most successful compression of the full Fourier bispectrum into a manageable number of measurements.

Compression by reduction to ≤ 30 bins

To this end we combine the power spectrum bins with a single additional configuration from the Fourier bispectrum or one of its proxies, and compute the corresponding Fisher matrix (as in Sec. 6.7.1) using values for $\partial\boldsymbol{\mu}/\partial\theta_\alpha$ estimated from our simulation suite. The four left panels of Fig. 6.15 show the reduction in predicted uncertainty — defined as the shrinkage of the error bar, $1-\sigma_{P+X}/\sigma_P$ — for the representative parameters σ_8 (solid lines) and w_0 (dotted lines) for each of the possible bins. Using these reductions as a measure of the information stored in each bin we conclude that most of the information carried by the Fourier bispectrum B is contained in small-scale triangles (towards larger triangle index). A similar conclusion applies for the line correlation function, for which significant reductions occur only for the first ~ 12 bins, corresponding to the range of scales $10\,h^{-1}\text{Mpc} - 50\,h^{-1}\text{Mpc}$. This is reasonable, because the line correlation is constructed to give a negligible signal on large scales. Finally, while the modal decomposition exhibits some variability, smaller mode numbers typically provide larger gains. The integrated bispectrum shows consistently weak improvements over all bins.

Second, for each combination $P+X$ we identify a set of 30 bins for X that provide the largest improvements. Adding them cumulatively to the power spectrum, starting from the bin carrying most information, we obtain the plot on the right-hand side of Fig. 6.15. Both the line correlation function and the modal bispectrum converge rapidly to the maximal improvement available from the entire set of bins that we measure (this is 30 bins for ℓ and 50 modes for β—see Table 6.2). For example, the line correlation is already within 2% of the maximum after we have added ~ 2 bins, while only ~ 5 modes of β are required to arrive at a similar value for the modal bispectrum. In comparison the Fourier bispectrum converges much more slowly to the maximum provided by the 95 bins that we measure. This is especially evident for σ_8, for which the improvement from the Fourier bispectrum has not yet converged to its maximum value after the 30^{th} bin.

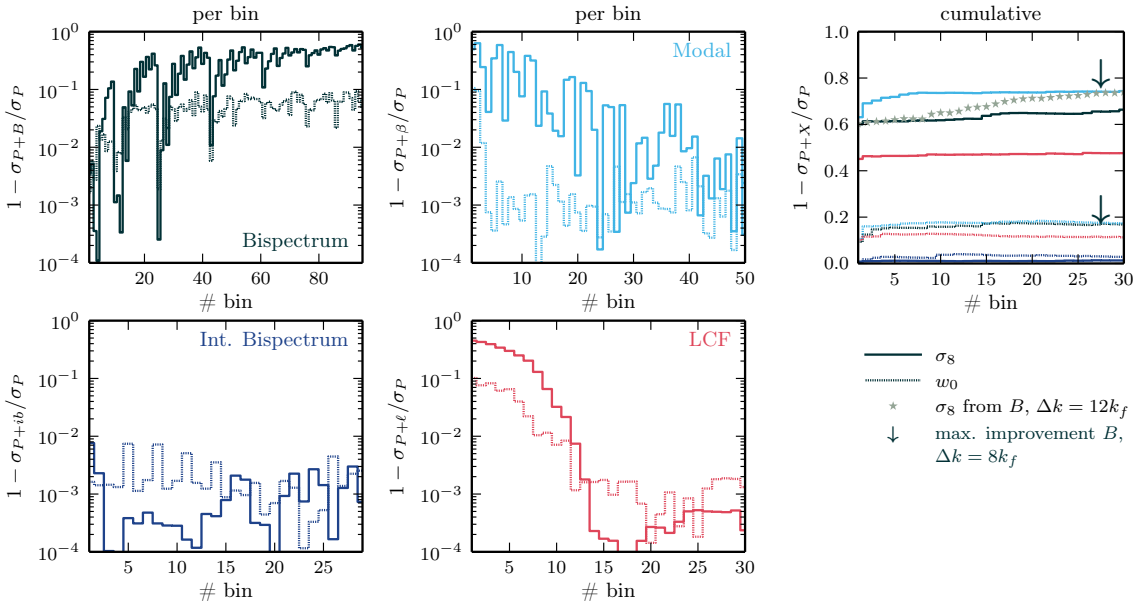


Fig. 6.15 *First two columns*: decrease in forecast parameter uncertainty (improvement) from combining the power spectrum with a single bin of a 3-point correlation measure, compared to the power spectrum alone. The Fisher matrix was computed from the non-Gaussian covariance matrix and the measured parameter derivatives $\partial\boldsymbol{\mu}/\partial\theta_\alpha$. Solid (dotted) lines show σ_8 (w_0) with all other parameters (including bias) marginalised. *Third column*: cumulative improvement from adding the 30 best bins. Arrows indicate the maximal improvement obtained from the Fourier bispectrum with $\Delta k = 8k_\mathrm{f}$, while stars show the uncertainty for σ_8 using Fourier bispectrum measurements with the larger bin width $\Delta k = 12k_\mathrm{f}$.

(For guidance, we mark this maximum value with black arrows on the plot.) However, it should be noted that our procedure to select the set of 30 bins is not optimal because it does not account for covariances between them. By analysing random subsets of the 95 possible bispectrum bins we find that faster convergence is possible, giving up to ~ 90% of the maximum reduction after 30 bins.

Compression by broadening bins

Rather than reducing the number of configurations by restriction to a subset, we might alternatively increase the width of each bin. The same volume of data would then be compressed into fewer measurements. To compare the performance of this strategy we repeat the analysis described above for the Fourier bispectrum with a broader bin width $\Delta k = 12k_\mathrm{f}$, which gives 34 rather than 95 Fourier configurations with $k_\mathrm{max} = 0.3\,h\,\mathrm{Mpc}^{-1}$. We plot the corresponding cumulative reduction in uncertainty for σ_8 as star-shaped symbols in the right-hand panel of Fig. 6.15. After 30 bins the improvement is similar to that obtained from the modal bispectrum, with the same caution about rate of convergence due to correlation between bins. Therefore—rather surprisingly—in this case we find no clear preference for the bin width $\Delta k = 8k_\mathrm{f}$ or $\Delta k = 12k_\mathrm{f}$, except that

$\Delta k = 8k_\mathrm{f}$ is more computationally expensive, and it is more difficult to find an optimal subset of configurations. However, it is not clear whether this conclusion would survive in a more realistic analysis, where the signal can be noisy and demands finer binning. To explore these issues in detail would require a more comprehensive analysis.

Results

This analysis agrees with the conclusions of Sections 6.7.3 and 6.7.4, and supports the modal bispectrum as a good choice of proxy for 3-point correlation data. In addition to the advantages discussed in previous sections, it requires the fewest bins and loses almost no information.

These results could be modified in cases where it is possible to compute a covariance matrix for $\gg 30$ configurations of the Fourier bispectrum, as done (for example) by Gil-Marín et al. (2017). However, the mock catalogues used to produce such covariance matrices are often generated using perturbation theory and therefore are likely to be inaccurate on small scales. We expect that it is a better strategy to use fewer bins and obtain high-quality measurements of the covariance matrix from catalogues generated using full N-body simulations. The significant benefit of the modal decomposition is that it facilitates construction of the smallest set of bins that still carry a majority of the information.

Finally, although the line correlation function provides weaker improvements than either the Fourier bispectrum or modal bispectrum, it has the advantage that it clearly separates the scales carrying useful information from those that do not—all bins with $r \gtrsim 50\,h^{-1}\mathrm{Mpc}$ have negligible impact. It is also possible that the performance of the line correlation function could be improved by relaxing the condition of strict collinearity, which would increase the range of Fourier configurations it is able to aggregate.

6.8.2 Shot Noise

Galaxies are discrete, point-like tracers of the underlying matter fluctuations, and therefore samples of their abundance are affected by shot noise. This noise is expected to impact higher-order statistics more significantly than the power spectrum (Chan and Blot, 2017; Sefusatti and Scoccimarro, 2005). Up to this point our analysis has implicitly used the low effective shot noise provided by our simulations, and therefore there is some concern that our forecasts will degrade with larger, more realistic noise. In this section we perform an approximate analysis of this degradation and quantify its effect on our predicted parameter uncertainties.

Assuming Poisson statistics, we may correct for shot-noise contributions to the observed discrete power spectrum $\hat{P}^\mathrm{disc}$ and bispectrum $\hat{B}^\mathrm{disc}$ by subtraction (Matarrese

et al., 1997; Peebles, 1980),

$$\hat{P}(k) = \hat{P}^{\rm disc}(k) - \frac{1}{\bar{n}}, \tag{6.70a}$$

$$\hat{B}(k_1,k_2,k_3) = \hat{B}^{\rm disc}(k_1,k_2,k_3) - \frac{1}{\bar{n}}\Big[\hat{P}(k_1)+\hat{P}(k_2)+\hat{P}(k_3)\Big] - \frac{1}{\bar{n}^2}. \tag{6.70b}$$

Here, $\bar{n}$ is the average number density of the discrete tracers. We use the upper and lower limits $\bar{n}_1 = 10^{-2}\,h^3\,{\rm Mpc}^{-3}$ and $\bar{n}_2 = 10^{-4}\,h^3\,{\rm Mpc}^{-3}$ to represent optimistic and pessimistic levels of shot noise for upcoming galaxy surveys. To measure $\hat{P}^{\rm disc}$ and $\hat{B}^{\rm disc}$ we downsample the number of particles in our simulation suite by selecting random subsets matching the desired averaged density $\bar{n}$, and use this to compute corrected estimators $\hat{P}$ and $\hat{B}$ from Eqs. (6.70a) and (6.70b). Although this downsampling procedure will not introduce exactly Poisson shot noise, we have checked that it is nearly Poisson by verifying that the corrected quantities agree with measurements made using the full set of particles to within a few percent. Strictly speaking, the covariance matrix of $\hat{P}$ and $\hat{B}$ obtained in this way is the matter covariance with Poisson shot noise, but for our fiducial biasing model we may interpret it as the covariance of the galaxy power spectrum and bispectrum with Poisson shot noise. We use this covariance, leaving the parameter derivatives unchanged from Sec. 6.7.2, to compute the Fisher matrices.

We plot forecasts using the fiducial number densities $\bar{n}_1$ and $\bar{n}_2$ in Fig. 6.16, with orange ellipses corresponding to the lower noise level (higher number density) and blue ellipses corresponding to the higher noise level (lower number density). The orange ellipses show good agreement with the forecasts for the idealised scenario of Sec. 6.7.2, indicating that relatively little degradation occurs. However, it is unlikely that such high number densities will be attained in the near future. By contrast the blue ellipses represent a conservative view of what should be possible.

If shot noise degrades the signal from 3-point correlations more strongly than for 2-point correlations then the fractional improvement from its inclusion should be smaller for low $\bar{n}$. In terms of Fig. 6.16 this means that the difference between the light and dark blue ellipses should be smaller than the difference between the light and dark orange ellipses. This effect is visible for some parameters, such as σ_8. However, in the case of Ω_m, w_0 and w_a the fractional improvement from inclusion of 3-point correlation data is *larger* at lower $\bar{n}$. The effect for w_0 and w_a is particularly striking. Using all particles in our simulations, the addition of B data decreased measurement uncertainties by 16% and 15%, respectively (see Table 6.3). With $\bar{n} = 10^{-4}\,h^3\,{\rm Mpc}^{-3}$ we find improvements of 41% and 36%. There are (at least) two possible explanations for this effect. First, while shot noise puts a stricter limit on the range of k-modes that are useful to recover cosmological information from 3-point correlations than from 2-point correlations, it also tends to render the respective covariance matrices more Gaussian, by relatively decreasing the correlation in the off-diagonal terms. That

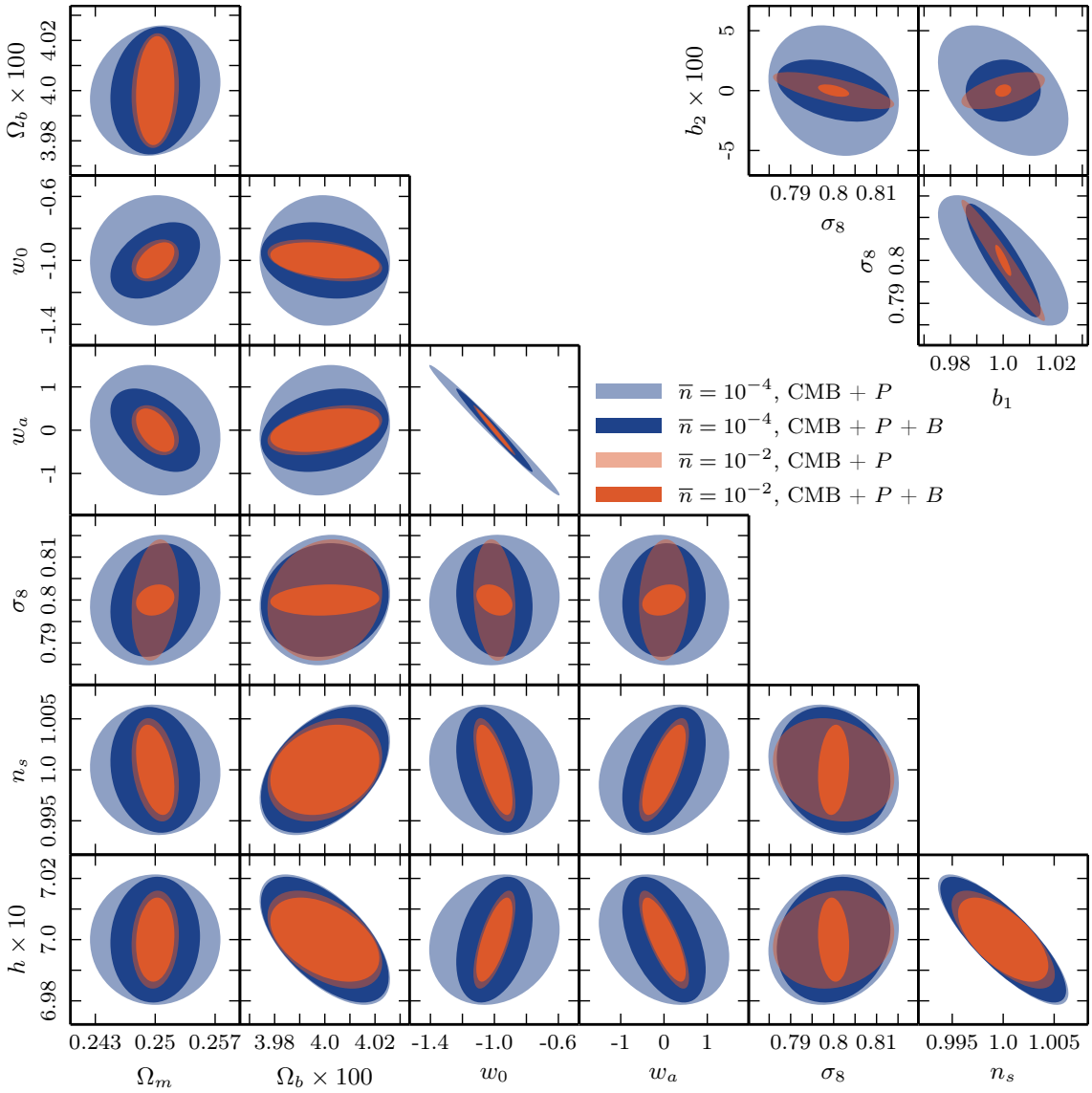


Fig. 6.16 Comparison of the Fisher forecasts with shot noise corresponding to $\bar{n}_1 = 10^{-2}\,h^3\,\mathrm{Mpc}^{-3}$ (orange) and $\bar{n}_2 = 10^{-4}\,h^3\,\mathrm{Mpc}^{-3}$ (blue). The pale ellipses correspond to uncertainties using the power spectrum only, while the dark ellipses show the predicted uncertainty when 3-point correlation information is included.

means different bins are less correlated and as the bispectrum allows for many more configurations than the power spectrum, this might enhance the relative importance of 3-point information. Alternatively, recovery of cosmological information in the presence of shot noise might significantly depend on cross-covariances between measurements. These cross-covariances themselves depend on the shot noise and can partially subtract its effect. By explicitly excluding the cross-covariance between $\hat{P}$ and $\hat{B}$, we find that the improvement in uncertainty from inclusion of the three point information drops slightly from 10% (9%) to 8% (7%) for w_0 (w_a) as shot noise increases, which strongly suggests that this second possibility is the driving force in the improvements of w_0 and w_a.

6.9 Conclusions

As large scale structure surveys grow in size and sophistication, the rapidly-approaching cosmic variance limit on 2-point statistics encourages us to look to higher-order correlations, such as the 3-point function, as a new source of information. Previously, Sefusatti et al. (2006) suggested that considerable additional constraining power could be achieved by combining the power spectrum and bispectrum. On the other hand, the signal-to-noise analysis in Chan and Blot (2017) concluded that the bispectrum yields no more than modest improvements. Our results show that there is a significant benefit from inclusion of three-point correlation data, but its benefits must be balanced against the challenges it brings.

In this chapter, we focus on two particular challenges: (1) The number of measurable configurations of the Fourier bispectrum is generally very large unless one coarse-grains the data. We have investigated whether the *modal bispectrum*, *line correlation function* and *integrated bispectrum* can act as 'proxies' for the Fourier bispectrum, compressing its information into fewer configurations without unacceptable information loss. (2) Bispectrum observations are difficult to model to the same accuracy as the power spectrum. Errors in clustering predictions from theoretical models, in addition to assumptions about covariances and noise properties, generally propagate into inaccurate error bars or a bias on inferred parameters. We have quantified how our forecasts are influenced by both the assumption of Gaussian covariance and theoretical errors.

To do so we have measured the power spectrum, Fourier bispectrum and each of its proxies from a suite of 200 dark matter N-body simulations at redshifts $z = 0$, $z = 0.52$ and $z = 1$ to obtain fully non-Gaussian covariances and cross-covariances. We measure the dependence of each measurement on the cosmological parameters $\{\Omega_m, \Omega_b, w_0, w_a, \sigma_8, n_s, h\}$ using additional simulations displaced from our fiducial model. We assume a local Lagrangian biasing scheme that includes two bias parameters, $\{b_1, b_2\}$. Using all these components, in combination with theoretical predictions for each proxy from tree-level and 1-loop SPT and the halo model, we have conducted a signal-to-noise analysis and implemented the Fisher forecasting method for an idealised survey scenario. Our main results on the constraining power and future viability of each measure of 3-point correlations are:

6.9.1 Comparison of 3-point correlation measures

Sec. 6.7.2 presented our main results. Our forecasts show that inclusion of the Fourier bispectrum offers significant improvements over the power spectrum alone, with $\mathcal{O}(10\% - 30\%)$ improvement on cosmological parameter constraints, and up to $\mathcal{O}(80\%)$ improvement when it is used to break degeneracies with the bias parameters. The *modal bispectrum* offers an attractive alternative, achieving equivalent constraints with

as few as 10 modes. However, up to 50 modes may be necessary to reconstruct the Fourier bispectrum to within $\lesssim 10\%$ accuracy on individual triangle configurations. The *line correlation function* appears to be slightly less optimal, although a future extension to sample more Fourier configurations by relaxing the requirement of strict collinearity may improve its performance. The *integrated bispectrum* offers little constraining power for our set of cosmological parameters. It is sensitive to highly squeezed triangles, whereas the gravitational bispectrum peaks on equilateral triangles. This property of ib is a disadvantage for our purposes, but may be an advantage if one is interested in studying squeezed-mode primordial non-Gaussianity with minimal degeneracies.

6.9.2 Data compression

In Sec. 6.8.1, we explored how the total constraining power of each measure is distributed over the total number of data bins. While the Fourier bispectrum and modal bispectrum give nearly equivalent parameter constraints when ~ 30 bins are used, the modal method converges to its full constraining power with a smaller subset of bins. We conclude that the modal bispectrum provides more efficient access to the information carried by 3-point correlations.

We note that more realistic survey scenarios—for example, accounting for noisy data—may require finer binning. Increasing the binning resolution of the Fourier bispectrum by a factor of n in each k-dimension corresponds to a factor $\mathcal{O}(n^3)$ increase in configurations. The number of simulations required to accurately capture their covariance would increase similarly. If the number of modal coefficients required to capture fine features of the bispectrum does not grow so dramatically, it is possible that the modal bispectrum could accumulate an even larger advantage compared to the Fourier bispectrum.

6.9.3 Signal-to-noise ratio as a measure of information content

In Sections 6.6, 6.7.2 and 6.7.5 we argued that use of the signal-to-noise ratio to predict the constraining power of 3-point correlation data can be misleading. We show that the bispectrum and phase bispectrum—which is probed by the modal bispectrum—give significantly different signal-to-noise ratios, but still yield nearly identical forecasts. As we describe in Sec. 6.7.5, for the scenarios considered in this chapter, the improvement shown by these forecasts is empirically better predicted by the signal-to-noise ratio of the phase bispectrum B_ϵ than the Fourier bispectrum B. The $\sim \mathcal{O}(30\%)$ uplift in signal-to-noise from the phase bispectrum translates to the same improvement in cosmological parameter constraints, except for those where degeneracies play a significant role. For a general parameter set and a given measure of the 3-point correlations, the signal-to-noise will not typically give an accurate estimate of its constraining power.

6.9.4 Impact of non-Gaussian covariances

Accounting for non-Gaussian covariance is essential for optimally constraining cosmological parameters. In Sec. 6.7.3 we showed that the Fourier bispectrum estimator is particularly sensitive to the covariance: our predicted uncertainties may be nearly a factor of 4 too small if the Gaussian approximation is used. At the same time, we find that the non-Gaussian cross-covariance between the power spectrum and the Fourier bispectrum or its proxies generally results in parameter errors that are $\mathcal{O}(10\%)$ *smaller* than if cross-covariances are ignored.

6.9.5 Impact of theoretical modelling uncertainties

Our results in Sec. 6.7.4 indicate that the impact of theory errors on our predicted uncertainties is smaller than the impact of assuming Gaussian covariance, although both approximations change the forecasts by ~ 30% to 50% on average. In this chapter we measure the effect of theoretical uncertainty by comparing forecasts using SPT and the halo model to forecasts derived purely from N-body measurements. Our approach differs from that of Baldauf et al. (2016) and Welling et al. (2016), who incorporated estimates of the theory error into their Fisher forecasts by taking the error in each data bin to be the sum of statistical and theoretical errors.

6.9.6 Impact of shot noise

To assess the impact of shot noise, in Sec. 6.8.2 we down-sample our simulation suite to averaged number densities of $\bar{n} = 10^{-2}\,h^3\,\mathrm{Mpc}^{-3}$ and $10^{-4}\,h^3\,\mathrm{Mpc}^{-3}$, and compute forecasts using non-Gaussian covariance matrices that include low and high levels of Poisson shot noise. Contrary to naïve expectations, we find that the addition of 3-point correlation information can become *more* significant at high levels of shot noise owing to the non-trivial dependence of the cross-covariance on $\bar{n}$. This appears most significant for the dark energy parameters w_0 and w_a, and suggests that 3-point correlation information may be crucial to distinguish between dark energy models. More generally, our result implies that 3-point correlation measurements may yield significant additional constraining power even when shot noise levels are high.

To make robust inferences with 3-point correlation information, future surveys will require refinement of the methods we have considered here. For example, while we have demonstrated that the modal decomposition provides efficient data compression of the matter bispectrum in an idealised survey, it will be important to verify that this remains true when halo distributions, redshift-space distortions and the complex noise properties of realistic surveys are introduced. We have emphasised the importance of including non-Gaussian covariances and theory uncertainties in our forecasts. Realistic

analyses will likely require more efficient ways to obtain covariances, and a consistent approach to inclusion of theory errors in software pipelines. Achieving each of these aims will be important milestones ahead of upcoming surveys of large-scale structure.

Appendices

Appendix 6.A Construction of the modal decomposition

6.A.1 Construction of the Q-basis

The goal of the modal decomposition is to write the estimated bispectrum in the form

$$w(k_1,k_2,k_3)\hat{B}(k_1,k_2,k_3) = \sum_{n=0}^{n_{\max}-1} \hat{\beta}_n^Q Q_n(k_1,k_2,k_3)\,, \tag{6.71}$$

where $w(k_1,k_2,k_3)$ is the arbitrary weighting function (6.24), and the Q_n represent basis modes with coefficients β_n^Q. The Q_n then contain all the information about the bispectrum. They should span the possible functions on wave numbers k_i that satisfy the triangle condition, $\sum_i k_i \geq 2\max\{k_1,k_2,k_3\}$ (denoted by $\mathcal{V}$ in the main text) but are otherwise arbitrary. For our concrete numerical results we choose a basis built out of one-dimensional polynomials $q_p(x)$ which are orthonormal within $\mathcal{V}$ (Fergusson et al., 2010). More precisely, in a unit box, we define the integral $\mathcal{T}[f] = \int_{\mathcal{V}} f(x)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$, where x,y,z satisfy the triangle condition within the box $x,y,z \in [0,1]$. Evaluating the y and z integrals, one finds that $\mathcal{T}[f] = 0.5\int_0^1 f(x)\,x(4-3x)\,\mathrm{d}x$. This allows one to define an inner product, $\langle f,g\rangle \equiv \mathcal{T}[fg]$ (which is not equal to the inner product (6.27)) and set up a generating function for the one-dimensional polynomials, q_n, using $w_n = \mathcal{T}[x^n]$, in the form of a secular determinant

$$q_n(x) = \frac{1}{\mathcal{N}} \begin{vmatrix} 1/2 & 7/24 & \dots & w_n \\ 7/24 & 1/5 & \dots & w_{n+1} \\ \dots & \dots & \dots & \dots \\ w_{n-1} & w_n & . & w_{2n-1} \\ 1 & x & \dots & x^n \end{vmatrix} \tag{6.72}$$

where $\mathcal{N}$ is chosen such that $\langle q_n, q_m\rangle = \mathbf{1}_{nm}$. The basis functions $Q_n(x,y,z)$ are defined as symmetric combinations of combinations of these 1-dimensional polynomials, in

the form

$$Q_n(x,y,z) = \frac{1}{6}\Big[q_r(x)q_s(y)q_t(z) + q_r(x)q_t(y)q_s(z) + \cdots + q_t(x)q_s(y)q_r(z)\Big]$$
$$\equiv q_{\{r}(x)q_s(y)q_{t\}}(z)\,, \tag{6.73}$$

with n representing the triple of indices $\{r,s,t\}$. After choosing an ordering of these triples we can exchange n for a simpler integer label. For a particular realisation with wave numbers in the range $k_{\min}$ and $k_{\max}$ we use the notation $Q_n(k_1,k_2,k_3)$ to represent $Q_n(x_1,x_2,x_3)$, where $x_i = (k_i - k_{\min})/(k_{\max} - k_{\min}) \in [0,1]$.

6.A.2 Calculation of the modal coefficients using the voxel method

In Sec. 6.4.7 we explained how Eq. (6.51) reduces estimation of the modal coefficients from simulation or data to a single 3-dimensional integral over a product of three Fourier transforms $\mathscr{M}_n(\boldsymbol{x})$. If the bispectrum is given analytically, however, we may instead use the simpler Eq. (6.27) and compute the inner product using a sum of volumes of all 'voxels' within a cubic grid with linear spacing along each axis (k_1,k_2,k_3).

To calculate the volume of each voxel we relabel the coordinates as (x,y,z), rescaled so that $0 \le x,y,z \le 1$. We associate each of the 8 possible vertices of the voxel with a value $p_1,\ldots,p_8$, given by the product of Q_m and wB (or Q_m and Q_n in the case of $\langle\!\langle Q_m|Q_n\rangle\!\rangle$) at that vertex. Finally, we define an interpolation function f by writing

$$f(x,y,z) = a_1 + a_2x + a_3y + a_4z + a_5xy + a_6xz + a_7yz + a_8xyz. \tag{6.74}$$

The coefficients a_i may be obtained analytically in terms of the p_i. We assign the volume of the voxel to be zero if fewer than four of its vertices satisfy the triangle condition, while if all 8 vertices satisfying the triangle condition its volume is

$$\int_{0\le x,y,z\le 1} f(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \frac{1}{8}\sum_{i=1}^{8} p_i, \tag{6.75}$$

as expected. For intermediate cases we write the volume in the form

$$\int_{\mathscr{C}} f(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z, \tag{6.76}$$

where $\mathscr{C}$ indicates that only those points satisfying the triangle condition and forming a closed volume within the voxel should be included. In the case of 4 points there are 3 possible volumes given by

$$\mathscr{C}_a^{(4)} = \{x,y,z \mid x+1 \le y+z\}\,, \quad \mathscr{C}_b^{(4)} = \{x,y,z \mid y+1 \le x+z\}\,, \quad \mathscr{C}_c^{(4)} = \{x,y,z \mid z+1 \le x+y\}\,. \tag{6.77}$$

For 5 points the only possibility is that $x+y+z \geq 2\max\{x,y,z\}$, while for 6 and 7 points there are again 3 possibilities, given respectively by,

$$\begin{aligned}
\mathscr{C}_a^{(6)} &= \{x,y,z \mid x \leq y+z, y \leq x+z\},\\
\mathscr{C}_b^{(6)} &= \{x,y,z \mid x \leq y+z, z \leq x+y\},\\
\mathscr{C}_c^{(6)} &= \{x,y,z \mid y \leq x+z, z \leq x+y\},\\
\mathscr{C}_a^{(7)} &= \{x,y,z \mid x \leq y+z\},\\
\mathscr{C}_b^{(7)} &= \{x,y,z \mid y \leq x+z\},\\
\mathscr{C}_c^{(7)} &= \{x,y,z \mid z \leq x+y\}.
\end{aligned} \tag{6.78}$$

In each case the analytic form of the integral in terms of the vertex values p_i can be calculated easily. Computation of each integral using this voxel method is highly accurate and efficient.

Contributions

The work presented in this chapter was published in Byun et al. (2017) and is the outcome of a collaborative project involving several people. My contributions include the development of numerical codes for the power spectrum, bispectrum and integrated bispectrum (in addition to the LCF) and the measurement of all these quantities along with their parameter derivatives. I have also computed the covariance matrices, the signal-to-noise as well as Fisher forecasts for all statistics (as a check, the Fisher forecasts were also independently computed by Joyce Byun), and contributed to the content of Sec. 6.7.5. While I did minor edits throughout the complete manuscript, I made the following major contributions to the write-up of the project: the parts on the LCF and galaxy bias in Secs. 6.2 and 6.3; the parts detailing the simulations, density field, and choice of bins in Sec. 6.4; the section on signal-to-noise (Sec. 6.6); the first two subsections in Sec. 6.7; and the complete Sec. 6.8 along with the analyses presented therein. Moreover, I produced all figures except Fig. 6.10 to 6.14.

Chapter 7

Bias loop corrections to the galaxy bispectrum in the multi-point propagator approach

After considering various alternative probes of the bispectrum in relation to the covariance matrix problematic, we now turn to another challenge: the theoretical modelling of the bispectrum in the moderately non-linear regime. In particular, this chapter focuses on the galaxy distribution in real space and aims to extend the tree-level results from Sec. 3.2.2 to one-loop order in perturbation theory. This step seeks to restore consistency with state-of-the-art models used in recent power spectrum analyses, which is crucial for their joint measurement in the future. As the one-loop bispectrum requires perturbations up to fourth order (see Sec. 3.2.1), we need to extend the bias expansion used in previous chapters. Under certain assumptions detailed in this chapter, we derive a complete basis for galaxy bias and study it in the framework of the multi-point propagator approach. This formalism allows us to easily solve the bias renormalisation problem and facilitates the computation of the bispectrum. Finally, we present results from fitting our model to data from mock galaxy catalogues.

7.1 Introduction

The interpretation of data gathered by large-scale structure (LSS) surveys requires an accurate understanding of galaxy bias — the relation between galaxies (or any other luminous tracer) and the underlying matter distribution. Although the formation of galaxies involves highly non-linear, small-scale processes, on sufficiently large scales it can be understood perturbatively, where the physics of galaxy formation is absorbed into a set of unknown bias parameters (see Desjacques et al., 2018, for a review). This set is by now well established up to the one-loop level for the galaxy power spectrum

(Chan et al., 2012; McDonald and Roy, 2009), and has been successfully applied in recent survey analyses (e.g. Grieb et al., 2017; Hou et al., 2018; Sánchez et al., 2017). It is still largely an open problem, though, which and how many bias parameters are required by a model of the galaxy bispectrum, in order to accurately describe data over a similar range of scales as the power spectrum.

Moreover, for theoretical predictions of galaxy clustering beyond leading order, a difficulty arises which is due to a mismatch of the bias parameters from the perturbative expansion, and those an observationalist would define through the measurement of correlation functions. In the perturbative model the galaxy density fluctuations are linearly proportional to the matter fluctuations on the largest scales (Kaiser, 1984), whereas on smaller scales a variety of non-linear terms are thought to appear. These terms involve powers of the matter density (Coles, 1993; Gaztanaga and Frieman, 1994), as well as the tidal field (Catelan et al., 2000; McDonald and Roy, 2009). More generally, Chan et al. (2012) and Mirbabayi et al. (2015) showed that in the absence of velocity bias (i.e. galaxies and matter are comoving) all these terms derive from Galilean invariants of the gravitational and velocity potentials. However, the non-linear contributions can affect the clustering statistics on large scales by a degree comparable to the linear term, which raises the question whether this bias expansion is a convergent series. This problem was first pointed out by (McDonald, 2006) and is usually dealt with by a procedure that renormalises the parameters of the bias expansion. While tractable for the power spectrum (McDonald and Roy, 2009), this procedure becomes increasingly complicated for higher-order statistics such as the bispectrum (Assassi et al., 2014).

The purpose of this chapter is twofold: we aim to examine galaxy bias in light of the multi-point propagator approach, which has previously been proven to be an efficient method of deriving non-linear corrections to the matter density (Bernardeau et al., 2010, 2008, 2012; Crocce et al., 2012). We will argue that when extended to galaxies, multi-point propagators not only facilitate the computation of observables, but also offer the most natural conception of galaxy bias and let us circumvent the problem of renormalisation. As an application, we then intend to use this formalism to compute the bispectrum including all relevant contributions from galaxy bias up to the one-loop level in perturbation theory and compare our results to Assassi et al. (2014).

As bispectrum analyses have so far been limited to a simplified, tree-level bias model (Gil-Marín et al., 2015, 2017; Saito et al., 2014), putting it on an equal footing with the power spectrum marks an important step towards a reliable joint analysis. This is guaranteed to unlock a considerable amount of additional information (Byun et al., 2017; Chan and Blot, 2017; Sefusatti et al., 2006), and should be of great interest for current and future data sets such as the *extended Baryon Oscillation Spectroscopic Survey* (*eBOSS*), *DESI*, and Euclid.

This chapter is organised as follows. In Sec. 7.2 we introduce our main idea of using appropriately defined, physical bias parameters in order to avoid the renormalisation problem. Moreover, we derive a complete set of Galilean invariant basis operators, that can be easily extended to arbitrarily high orders. We develop the multi-point propagator formalism in Sec. 7.3 and compute the first few galaxy propagators based on the previously defined basis. These are used in Sec. 7.4 to compute the one-loop power spectrum and bispectrum, where we also discuss corrections due to the effective field theory (EFT) and shot noise. In Sec. 7.5 we apply the full one-loop bispectrum model for the first time to measurements from galaxy mock catalogues.

7.2 Galaxies as tracers of the dark matter field

7.2.1 Clustering statistics

We are interested in the statistical properties of the observed galaxy distribution. These are commonly quantified by a hierarchy of correlation functions of the density perturbations, which we write as $\delta_g(\boldsymbol{x}) = \rho_g(\boldsymbol{x})/\bar{\rho}_g - 1$. The subscript 'g' here stands for galaxies, but it could equally well denote any other tracer of the mass field, such as quasars, galaxy clusters or the Ly-α forest. The two- and three-point correlation functions in Fourier space — the power spectrum and bispectrum — are defined, respectively, as

$$\langle \delta_g(\boldsymbol{k})\, \delta_g(\boldsymbol{k}') \rangle \equiv (2\pi)^3\, P_g(k)\, \delta_D(\boldsymbol{k}+\boldsymbol{k}')\,, \tag{7.1}$$

$$\langle \delta_g(\boldsymbol{k}_1)\, \delta_g(\boldsymbol{k}_2)\, \delta_g(\boldsymbol{k}_3) \rangle \equiv (2\pi)^3\, B_g(k_1,k_2,k_3)\, \delta_D(\boldsymbol{k}_{123})\,, \tag{7.2}$$

where $\boldsymbol{k}_{123} \equiv \boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3$ and the appearance of the delta distribution is a manifestation of statistical homogeneity. Statistical isotropy further demands that the power spectrum only depends on the magnitude of the two wave vectors participating in the correlator, while the bispectrum is a function of three wave numbers k_1, k_2 and k_3. Finally, to pass from configuration to Fourier space we have adopted the convention

$$\delta_g(\boldsymbol{x}) = \int_{\boldsymbol{k}} \exp\left(-i\,\boldsymbol{k}\cdot\boldsymbol{x}\right) \delta_g(\boldsymbol{k})\,, \tag{7.3}$$

using a short-hand notation for k-space integrals, i.e. $\int_{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n} \equiv \int \mathrm{d}^3k_1/(2\pi)^3 \cdots \mathrm{d}^3k_n/(2\pi)^3$.

Analogous definitions hold for the dark matter field, which, as the dominant matter component in our Universe, heavily influences the clustering of galaxies. At any given time τ it is fully characterised by its density perturbations $\delta(\boldsymbol{x},\tau)$ and peculiar velocity $\boldsymbol{v}(\boldsymbol{x},\tau)$. On large scales it is furthermore an excellent approximation of taking the velocity field to be irrotational, in which case it is given by a single scalar degree of freedom — the velocity divergence $\theta(\boldsymbol{x},\tau) \equiv \nabla\cdot\boldsymbol{v}(\boldsymbol{x},\tau)$. These characteristic quantities evolve

under the influence of the gravitational potential, sourced by the matter perturbations themselves, and the equations of motion for δ and θ follow from imposing conservation of mass and momentum (see Bernardeau et al., 2002, for a review).

In standard perturbation theory (SPT) these equations are solved as series expansions about the linear density field $\delta^{(1)} \equiv \delta_L$, such that

$$\delta(\boldsymbol{x},\tau) = \sum_{n=1}^{\infty} D^n(\tau)\,\delta^{(n)}(\boldsymbol{x})\,, \tag{7.4}$$

where, to very good accuracy, all cosmology dependence is encoded in the linear growth factor $D(\tau)$ (Bouchet et al., 1992, 1995). In Fourier space the n-th order solution is constructed out of n powers of the linear density field, which are coupled via the SPT kernels F_n:

$$\delta^{(n)}(\boldsymbol{k}) = (2\pi)^3 \int_{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n} \delta_D(\boldsymbol{k}-\boldsymbol{k}_{1\cdots n})\,F_n(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n)\,\delta_L(\boldsymbol{k}_1)\cdots\delta_L(\boldsymbol{k}_n)\,. \tag{7.5}$$

The velocity divergence can be expanded in a similar manner and the n-th order solutions are obtained by replacing F_n with G_n in the equation above. Explicit expressions for these kernels can be found in Bernardeau et al. (2002). We also note that at linear order, we have $\delta = \theta$.

Throughout this chapter we will assume that the linear dark matter density field is Gaussian, implying that all of its statistical information is contained in the power spectrum, which we denote by $\langle\delta_L(\boldsymbol{k})\,\delta_L(\boldsymbol{k}')\rangle = (2\pi)^3\,P_L(k)\,\delta_D(\boldsymbol{k}+\boldsymbol{k}')$. However, because of the mode coupling in the non-linear regime, gravitational evolution produces a non-zero bispectrum as well. At leading or tree-level order the bispectrum is given by a correlation of a second order and two linear modes, $\langle\delta^{(2)}\delta^{(1)}\delta^{(1)}\rangle$, and can be expressed by the linear power spectrum as follows (Bernardeau et al., 2002; Fry, 1984):

$$B^{\rm tree}(k_1,k_2,k_3) = 2\,F_2(\boldsymbol{k}_1,\boldsymbol{k}_2)\,P_L(k_1)\,P_L(k_2) + \text{cyc.} \tag{7.6}$$

Next-to-leading order or one-loop corrections for both the matter power spectrum and bispectrum are build from consecutively higher SPT solutions. Whilst the former includes the terms $\langle\delta^{(2)}\delta^{(2)}\rangle$ and $\langle\delta^{(3)}\delta^{(1)}\rangle$, and therefore involves third order SPT, for the 1-loop bispectrum we need to consider the density contrast up to fourth order. These corrections have been well studied in the literature and were first derived in Jain and Bertschinger (1994); Juszkiewicz (1981); Scoccimarro (1997); Scoccimarro et al. (1998a); Suto and Sasaki (1991); Vishniac (1983).

7.2.2 Basics of the (renormalised) bias expansion

In order to compute the galaxy power spectrum and bispectrum, we require a relation between the galaxy and matter perturbations, which is usually written as some function $\delta_g(\delta_L)$ which is then Taylor expanded. By illustrating the issues one faces when calculating galaxy correlators beyond leading order with such an approach, we introduce an alternative bias expansion in this section, which differs from the usual one in the sense that its coefficients are measurable and thus *physical* quantities. The measurable quantities are cross-correlations of the galaxy field with the matter field, which corresponds to the so-called propagators in RPT and its generalisations Bernardeau et al. (2008); Crocce and Scoccimarro (2006b). In fact, the only difference here is that the non-linear mapping we are interested in maps matter to galaxies, instead of linear to non-linear matter perturbations. So, in this respect there is nothing new, just that the SPT kernels are replaced by bias kernels, as we detail below.

Let us start with the simplest and most well-known relation, that of local bias, where the galaxy fluctuations are considered to be a local function of the matter fluctuations and expanded in a Taylor series around $\delta = 0$ (Coles, 1993; Fry and Gaztanaga, 1993). Dropping the position argument from all density perturbations, we have

$$\delta_g = \sum_n \frac{1}{n!}\left(\frac{\partial^n \delta_g}{\partial \delta^n}\right)_0 \delta^n \equiv \bar{b}_0 + \bar{b}_1\,\delta + \frac{\bar{b}_2}{2}\delta^2 + \frac{\bar{b}_3}{3!}\delta^3 + \ldots \tag{7.7}$$

where $\bar{b}_n = (\partial^n \delta_g/\partial\delta^n)_0$ $(n>0)$ with $()_0$ denoting evaluation at $\delta = 0$, and $\bar{b}_0$ enforces $\langle\delta_g\rangle = 0$. The bias parameters so defined are not observables as we only measure correlators (expectation values or ensemble averages of fields), not quantities evaluated at $\delta = 0$. This presents an issue that becomes immediately apparent, if we consider the galaxy power spectrum up to one-loop order (i.e. the first "bias loop"). Plugging Eq. (7.7) into Eq. (7.1) we obtain

$$P_g(k) = \bar{b}_1^2\,P_L(k) + \left[\bar{b}_1\bar{b}_3\sigma^2\right]P_L(k) + \frac{\bar{b}_2^2}{2}\int_{\boldsymbol{q}} P_L(|\boldsymbol{k}-\boldsymbol{q}|)P_L(q) + \ldots \tag{7.8}$$

where $\sigma^2 \equiv \langle\delta^2\rangle$, the dots denote two-loop and higher contributions, and for simplicity we assumed that the bias expansion is done in Lagrangian space so that matter fluctuations are Gaussian (this corresponds to setting all SPT kernels F_n to zero). We notice that the first term of the one-loop contribution in square brackets is proportional to $P_L(k)$ and can therefore have a strong impact on the large-scale galaxy power spectrum. However, we can redefine (or "renormalise", McDonald, 2006) the linear bias parameter to be $b_1 \equiv \bar{b}_1 + \bar{b}_3\,\sigma^2/2$, such that we retain the form

$$P_g(k) = b_1^2\,P_L(k) + \frac{\bar{b}_2^2}{2}\int_{\boldsymbol{q}} P_L(|\boldsymbol{k}-\boldsymbol{q}|)P_L(q) \tag{7.9}$$

Clearly, as more bias loops are included, the expression for b_1 keeps changing, but the principle remains the same — the observed linear bias is defined as (the square root of) the coefficient in front of P_L. A similar situation holds for for the quadratic bias parameter b_2. To see this consider now the bispectrum to one-loop, under the same assumptions (linear matter fluctuations) we have:

$$B_{123} = \bar{b}_2\bar{b}_1^2 P_1P_2 + \left[\bar{b}_1\bar{b}_2\bar{b}_3\sigma^2 + \frac{1}{2}\bar{b}_1^2\bar{b}_4\sigma^2\right]P_1P_2 + \frac{1}{2}\bar{b}_1\bar{b}_2\bar{b}_3\,P_1\int_{\boldsymbol{q}} P_L(|\boldsymbol{k}_2-\boldsymbol{q}|)P_L(q)$$
$$+\,\bar{b}_2^3\int_{\boldsymbol{q}} P_L(|\boldsymbol{k}_1-\boldsymbol{q}|)P_L(|\boldsymbol{k}_2+\boldsymbol{q}|)P_L(q) + \text{cyc.} + \ldots \tag{7.10}$$

where $P_i \equiv P_L(k_i)$ and cyc. denotes cyclic permutations of each term over the three wavevectors. Again, the first term in square brackets corresponds to the renormalisation of the linear bias seen in the power spectrum but in addition now there is also a second term that corresponds to a renormalisation of the *quadratic bias* $b_2 \equiv \bar{b}_2 + \bar{b}_4\sigma^2/2$, so that we can write

$$B_{123} = b_2 b_1^2 P_1P_2 + \frac{1}{2}\bar{b}_1\bar{b}_2\bar{b}_3\,P_1\int_{\boldsymbol{q}} P_L(|\boldsymbol{k}_2-\boldsymbol{q}|)P_L(q)$$
$$+\,\bar{b}_2^3\int_{\boldsymbol{q}} P_L(|\boldsymbol{k}_1-\boldsymbol{q}|)P_L(|\boldsymbol{k}_2+\boldsymbol{q}|)P_L(q) + \text{cyc.} \tag{7.11}$$

A number of questions arise from this procedure. 1) Is the b_2 renormalisation that follows from the one-loop bispectrum consistent with the one that follows from the two-loop power spectrum? If so, is that true for all other bias parameters? 2) Is there a way to do all these renormalisations 'automatically', instead of calculating statistic by statistic, and order by order? 3) Is there a simple connection between renormalisations of different N-point correlators?

The answer to all these questions is 'yes' (Bernardeau et al., 2008, 2012; Crocce and Scoccimarro, 2006a,b). Since the terms that renormalise bias factors are factorisable loop corrections (i.e. they take the form of lower-loop corrections multiplied by new contributions — in the example above functions of bias and σ^2), they correspond to diagrams which are called *reducible*. That means they can be decomposed into two or more connected diagrams by cutting one or more internal lines. In general, any loop correction to a given statistic can be classified into either irreducible (cannot be decomposed into connected diagrams by cutting an internal line) and reducible (the diagrams that describe the renormalisation procedure above).

Therefore it is desirable to construct perturbation theories (PT) in terms of the sum over all (to include renormalisations to arbitrary loop order) reducible diagrams with a given number of external lines, instead of the usual SPT kernels. Such objects are known as multi-point propagators (see Fig. 7.1), and correspond to the expectation

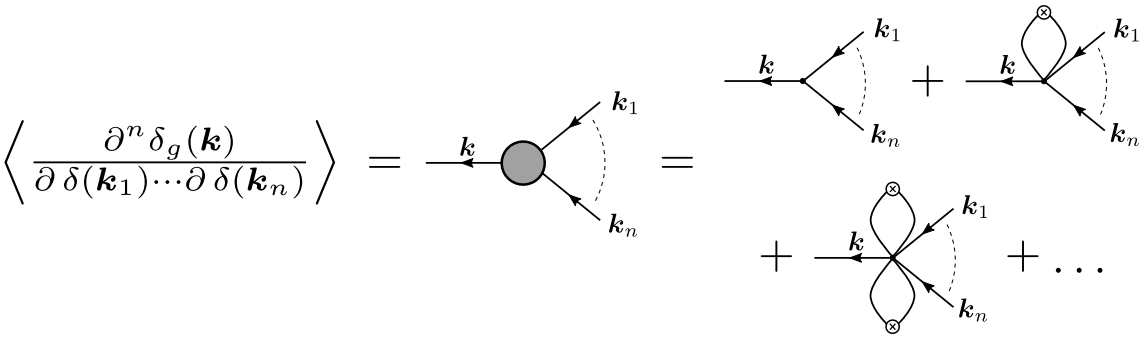


Fig. 7.1 The n-th multi-point propagator is given by the sum of all reducible diagrams with n legs of incoming momentum $\boldsymbol{k}_1, \ldots, \boldsymbol{k}_n$. The vertices correspond to the kernels associated with terms in the bias expansion, e.g. the vertices of the three diagrams on the right-hand side are given by b_n, b_{n+2} and b_{n+4} for the expansion in Eq. (7.7). If terms non-local in the matter density enter the bias relation, these kernels acquire a scale dependence based on all incoming momenta. The crossed circles stand for power spectra.

value of functional derivatives of the galaxy field with respect to the matter field: the derivatives produce the external lines, while the expectation value generates the loops. In addition, the fact that a given multi-point propagator appears in different statistics, establishes the connection between renormalisations of different N-point functions; e.g. the renormalisation of b_2 that follows from the one-loop bispectrum is consistent with the one that follows from the two-loop power spectrum because in both cases they are incorporated into the two-point propagator. As a result of this, $\bar{b}_2$ in Eq. (7.9) can be replaced by b_2 and similarly $\bar{b}_1\bar{b}_2\bar{b}_3 \to b_1 b_2 b_3$ and $\bar{b}_2^3 \to b_2^3$ in Eq. (7.11).

Therefore, the physical bias parameters of order n that appear in correlators correspond to the sum over all reducible diagrams with n external legs. For example when $n = 1$, i.e. linear bias, we simply have:

$$b_1 = \left\langle \frac{\partial \delta_g}{\partial \delta} \right\rangle = \bar{b}_1 + \bar{b}_3 \frac{\sigma^2}{2} + \ldots = \sum_{n=0}^{\infty} \frac{\bar{b}_{2n+1}}{n!} \left(\frac{\sigma^2}{2}\right)^n , \tag{7.12}$$

which corresponds to the sum of all reducible diagrams with a single external leg (corresponding to P_L at k). Similarly, for $n = 2$ we have from Eq. (7.7)

$$b_2 = \left\langle \frac{\partial^2 \delta_g}{\partial \delta^2} \right\rangle = \bar{b}_2 + \bar{b}_4 \frac{\sigma^2}{2} + \ldots = \sum_{n=0}^{\infty} \frac{\bar{b}_{2n+2}}{n!} \left(\frac{\sigma^2}{2}\right)^n , \tag{7.13}$$

that is the sum over all reducible diagrams with two external legs (corresponding to P_L at k_1 and k_2 as seen in the leading order bispectrum, Eq. 7.11). Clearly, the calculations in Eqs. (7.12-7.13) are significantly easier than performing the renormalisation procedure leading to Eqs. (7.9,7.11) order by order and for each statistic individually.

Another crucial property of the multi-point propagators is that for a Gaussian δ, they can be shown to be the cross-correlation bias between galaxies and matter fluctuations (Crocce and Scoccimarro, 2006a), and for the linear and quadratic bias cases we have

$$\langle \delta_g \, \delta \rangle = \left\langle \frac{\partial \delta_g}{\partial \delta} \right\rangle \langle \delta \, \delta \rangle \,, \tag{7.14}$$

$$\langle \delta_g \, \delta \, \delta \rangle = \left\langle \frac{\partial^2 \delta_g}{\partial \delta^2} \right\rangle \langle \delta \, \delta \rangle \, \langle \delta \, \delta \rangle \,. \tag{7.15}$$

That is, the observables corresponding to these objects are no other than the standard cross-correlation Lagrangian bias coefficients routinely measured in N-body simulations, see e.g. Bel et al. (2015); Pollack et al. (2012); Saito et al. (2014); Tinker et al. (2010).

Turning the argument around, this means a physical bias expansion should be constructed out of the sum of reducible diagrams with a given number of external legs, in effect, trading the kernels for multi-point propagators:

$$\left(\frac{\partial^n \delta_g}{\partial \delta^n} \right)_0 \longrightarrow \left\langle \frac{\partial^n \delta_g}{\partial \delta^n} \right\rangle . \tag{7.16}$$

This allows us to remove the disconnect between the parameters appearing in the standard bias expansion, Eq. (7.7), and those in correlators, Eqs. (7.9,7.11). We obtain a new expansion of the form (Bernardeau et al., 2008, 2012):

$$\begin{aligned} \delta_g &= \left\langle \frac{\partial \delta_g}{\partial \delta} \right\rangle \delta + \frac{1}{2!} \left\langle \frac{\partial^2 \delta_g}{\partial \delta^2} \right\rangle \left[\delta^2 - \langle \delta^2 \rangle \right] + \frac{1}{3!} \left\langle \frac{\partial^3 \delta_g}{\partial \delta^3} \right\rangle \left[\delta^3 - 3 \langle \delta^2 \rangle \delta - \langle \delta^3 \rangle \right] + \ldots \\ &= b_1 \delta + \frac{b_2}{2!} \left[\delta^2 - \langle \delta^2 \rangle \right] + \frac{b_3}{3!} \left[\delta^3 - 3 \langle \delta^2 \rangle \delta - \langle \delta^3 \rangle \right] + \ldots \end{aligned} \tag{7.17}$$

where the structure of the square brackets is given by δ^n minus all possible actions of $\langle\rangle$ on δ^n with a constant term that respects that the expectation value is zero for non-Gaussian δ, and the second equality assumes local bias as we have done so far. If the expansion in Eq. (7.17) is done with respect to a Gaussian δ, for instance the linear density fluctuations, terms such as $\langle \delta^3 \rangle$ and $\langle \delta^4 \rangle_c$ will vanish and we will be expanding δ_g in terms of Hermite polynomials, as was already suggested in Szalay (1988). Equation (7.17) automatically satisfies $\langle \delta_g \rangle = 0$ and indeed takes into account all renormalisations, thus replacing the propagators by numbers gives precisely the renormalised local bias expansion. We stress that in this bias expansion the b_i are renormalised bias parameters and they correspond to the cross-correlation (Lagrangian) bias coefficients measured in simulations. Therefore, we could have started with Eq. (7.17) as our expansion instead of Eq. (7.7) and we would have never had to talk about renormalisation of the bias parameters.

Note that for the local bias expansion discussed so far, the functional derivatives in Eqs. (7.12-7.13) are most easily computed in configuration space and result in simple numbers. However, when considering a more general bias scheme (as we will do in Sec. 7.2.3), the multi-point propagators become functions of position, or functions of momenta $\boldsymbol{k}_i$ when expressed in Fourier space:

$$\left\langle \frac{\partial^2 \delta_g}{\partial \delta^2} \right\rangle \longrightarrow \left\langle \frac{\partial^2 \delta_g(\boldsymbol{k})}{\partial \delta(\boldsymbol{k}_1)\, \partial \delta(\boldsymbol{k}_2)} \right\rangle, \tag{7.18}$$

as was already shown in Fig. 7.1.

7.2.3 A Galilean invariant basis for local galaxy bias

In this section we aim to provide a complete basis for the general bias expansion, required by the bispectrum at one-loop order. By basis we mean a set of linearly independent operators at each order of perturbation theory. We will largely follow up on the earlier work of Chan et al. (2012); Desjacques et al. (2018); Mirbabayi et al. (2015), but distinguish between bias relations when tracers are either identified at the level of the initial conditions ("Lagrangian") or from the late-time, non-linear density field ("Eulerian"). Our choice of basis will make this distinction explicit and therefore differs in the type of operators from that given in Desjacques et al. (2018); Mirbabayi et al. (2015).

Galileons as the general basis operators

Let us begin with two physical scales important for the process of galaxy formation: 1) the spatial extent R_* on which this process depends on the precise distribution of matter, and 2) the typical time T_* it takes for this matter distribution to collapse into a bound object. While the latter is a significant fraction of the Hubble time H^{-1}, the scale R_* usually corresponds to the Lagrangian radius of the galaxy's host halo, which is of the order ~ 1 Mpc. If we are interested in the clustering of galaxies on scales $r \gg R_*$, then we can consider galaxy formation as essentially local in space. For now we will take this to be the case, before relaxing this assumption in Sec. 7.2.4.

A simple conclusion from the long timescales associated with galaxy formation is that the galaxy density is unlikely to depend solely on the matter density at the same time and position, as assumed by the expansion considered in Sec. 7.2.2. Consequently, other properties of the matter field such as the tidal field must enter the bias relation (Catelan et al., 2000; McDonald and Roy, 2009). More generally, we can argue that it

should depend on the gravitational and velocity potentials, defined by

$$\nabla^2 \Phi(\boldsymbol{x},\tau) = \delta(\boldsymbol{x},\tau)\,, \tag{7.19}$$

$$\nabla^2 \Phi_v(\boldsymbol{x},\tau) = \theta(\boldsymbol{x},\tau)\,, \tag{7.20}$$

as these drive the time evolution in the regime where the dark matter flow is irrotational. According to the equivalence principle all leading local gravitational effects must stem from second derivatives, which we write as $\nabla_{ij}\Phi(\boldsymbol{x},\tau) \equiv \partial_i\partial_j\Phi(\boldsymbol{x},\tau)$. Similarly, if we assume that dark matter and galaxies are comoving (i.e. no velocity bias), Galilean invariance of the equations of motion (Scoccimarro and Frieman, 1996) implies that only second derivatives of the velocity potential are allowed to appear.

Furthermore, as δ_g is a scalar and therefore invariant under spatial coordinate transformations, we can limit ourselves to all scalar invariants of the tensors $\nabla_{ij}\Phi$ and $\nabla_{ij}\Phi_v$. In three dimensions the Cayley-Hamilton theorem guarantees that there can only be three such invariants, which can be expressed by the so-called *Galileons* (Chan et al., 2012) (repeated indices are summed over):

$$\mathscr{G}_1(\Phi) \equiv \nabla^2\Phi\,, \tag{7.21}$$

$$\mathscr{G}_2(\Phi) \equiv \left(\nabla_{ij}\Phi\right)^2 - \left(\nabla^2\Phi\right)^2\,, \tag{7.22}$$

$$\mathscr{G}_3(\Phi) \equiv \left(\nabla^2\Phi\right)^3 + 2\nabla_{ij}\Phi\,\nabla_{jk}\Phi\,\nabla_{ki}\Phi - 3\left(\nabla_{ij}\Phi\right)^2\nabla^2\Phi\,, \tag{7.23}$$

and similarly for Φ_v. We note that their leading SPT expressions are of first, second, and third order, respectively. By inverting the Poisson equation (Eq. 7.19) we can derive their Fourier space analogues and for the latter two we obtain:

$$\mathscr{G}_2(\boldsymbol{k}|\Phi) = \int_{\boldsymbol{k}_1,\boldsymbol{k}_2} K(\boldsymbol{k}_1,\boldsymbol{k}_2)\,\delta(\boldsymbol{k}_1)\,\delta(\boldsymbol{k}_2)\,\delta_D(\boldsymbol{k}-\boldsymbol{k}_{12})\,, \tag{7.24}$$

$$\mathscr{G}_3(\boldsymbol{k}|\Phi) = \int_{\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3} L(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)\,\delta(\boldsymbol{k}_1)\,\delta(\boldsymbol{k}_2)\,\delta(\boldsymbol{k}_3)\,\delta_D(\boldsymbol{k}-\boldsymbol{k}_{123})\,, \tag{7.25}$$

where we have defined the following two kernel functions,

$$K(\boldsymbol{k}_1,\boldsymbol{k}_2) \equiv \mu_{12}^2 - 1\,, \tag{7.26}$$

$$L(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) \equiv 2\,\mu_{12}\,\mu_{23}\,\mu_{31} - \mu_{12}^2 - \mu_{23}^2 - \mu_{31}^2 + 1\,, \tag{7.27}$$

with $\mu_{ij} \equiv \boldsymbol{k}_i\cdot\boldsymbol{k}_j\,/\,k_i\,k_j$. Similar expressions hold for Φ_v by replacing δ with θ in Eqs. (7.24) and (7.25).

Lagrangian basis

Let us now consider the bias relation on some initial time slice. In that case we are only dealing with linear quantities and the single degree of freedom is Φ_L, as at linear order we have $\Phi = \Phi_v$. Consequently, the only terms that can appear in the bias relation for objects at the initial conditions are $\mathcal{G}_n(\Phi_L)$, such that there will be n basis operators at n-th order in the expansion, i.e.

$$\begin{aligned}
&\text{1st:} \quad \mathcal{G}_1(\Phi_L)\,,\\
&\text{2nd:} \quad \mathcal{G}_1(\Phi_L)^2\,,\ \mathcal{G}_2(\Phi_L)\,,\\
&\text{3rd:} \quad \mathcal{G}_1(\Phi_L)^3\,,\ \mathcal{G}_1(\Phi_L)\mathcal{G}_2(\Phi_L)\,,\ \mathcal{G}_3(\Phi_L)\,,\\
&\text{4th:} \quad \mathcal{G}_1(\Phi_L)^4\,,\ \mathcal{G}_1(\Phi_L)^2\mathcal{G}_2(\Phi_L)\,,\ \mathcal{G}_1(\Phi_L)\mathcal{G}_3(\Phi_L)\,,\ \mathcal{G}_2(\Phi_L)^2\,,
\end{aligned} \tag{7.28}$$

and we assign a free bias parameter to each of these operators. We stress that the corresponding tracer density at initial time will thus be a local function of $\nabla_{ij}\Phi_L$, which is similar in spirit with more phenomenological approaches, such as the peak and excursion set bias models — in both cases the tracer density is defined via local procedures on $\nabla_{ij}\Phi_L$.

At late times, where we actually observe and identify galaxies, the gravitational and velocity potentials are clearly different, which implies that the basis in Eq. (7.28) will be insufficient. However, the operators that are missing in Eq. (7.28) will all be generated through time evolution. That is, even though we did not allow for certain operators at the initial time, they will automatically be included in the evolved tracer density, but with fixed coefficients that are functions of the remaining bias parameters (Chan et al., 2012).

Eulerian basis

To complete the basis in Eq. (7.28) it seems obvious to simply double the number of operators at each order by including a set of Galileons for both, the evolved Φ and Φ_v, and also allow for their combinations. Unfortunately, this produces a lot of redundancy as many of these operators are degenerate, so our task will be to identify those, which are linearly independent. We follow the strategy first developed in Chan et al. (2012).

At first order in SPT we have already established that $\mathcal{G}_1(\Phi) = \mathcal{G}_1(\Phi_v)$, and we choose the former, i.e. the matter fluctuation itself, as our first basis operator. Likewise, the two second-order Galileons are degenerate at second order in SPT and furthermore, we have

$$\mathcal{G}_1^{(2)}(\Phi) - \mathcal{G}_1^{(2)}(\Phi_v) = \delta^{(2)}(\boldsymbol{x}) - \theta^{(2)}(\boldsymbol{x}) = -\frac{2}{7}\mathcal{G}_2(\Phi_L)\,, \tag{7.29}$$

proving that the basis in Eq. (7.28) is complete up to that order. The need for an additional operator occurs for the first time at third order. Using the notation $\Delta_n \mathscr{G}_m \equiv \mathscr{G}_m^{(n)}(\Phi) - \mathscr{G}_m^{(n)}(\Phi_v)$ for the difference between the m-th Galileons evaluated at n-th order in SPT, we see that in addition to the ones already written in Eq. (7.28) there are the following four combinations:

$$\Delta_3 \mathscr{G}_1\,,\ \Delta_3 \mathscr{G}_2\,,\ \delta\,\Delta_2 \mathscr{G}_1\,,\ \Delta_3 \mathscr{G}_1^2\,. \tag{7.30}$$

From Eq. (7.29) follows that the latter two are degenerate with $\delta\,\mathscr{G}_2(\Phi_v)$, while $\Delta_3\mathscr{G}_1$ contains a contribution that cannot be written in terms of second derivatives of Φ_v and is thus not Galilean invariant. This only leaves the second combination, which gives

$$\Delta_3 \mathscr{G}_2 = -\frac{4}{7}\Big[\nabla_{ij}\Phi_L \nabla_{ij}\nabla^{-2}\mathscr{G}_2(\Phi_L) - \delta\,\mathscr{G}_2(\Phi_L)\Big]\,, \tag{7.31}$$

demonstrating that the additional basis operators induced by gravity can no longer be expressed as local functions of the linear gravitational potential, i.e. $\nabla_{ij}\Phi_L$. This is a well-known fact, gravitational instability is non-local and this first manifests itself in the (second-order) corrections to the Zel'dovich approximation (Kofman and Pogosyan, 1995).

Instead of explicitly calculating the differences between Galileons of Φ and Φ_v, we can follow an alternative strategy (which builds on Chan et al. (2012)) that will prove particularly useful for extending the basis beyond third order. For that purpose let us consider Lagrangian perturbation theory (LPT), which summarises all of the dynamics in its Lagrangian displacement field

$$\boldsymbol{\Psi}(\boldsymbol{q},\tau) = D_1(\tau)\boldsymbol{\Psi}^{(1)}(\boldsymbol{q}) + D_2(\tau)\boldsymbol{\Psi}^{(2)}(\boldsymbol{q}) + \dots\,, \tag{7.32}$$

that moves particles from their initial positions $\boldsymbol{q}$ to their final destinations $\boldsymbol{x} = \boldsymbol{q} + \boldsymbol{\Psi}(\boldsymbol{q},\tau)$. The functions $\boldsymbol{\Psi}^{(n)}(\boldsymbol{q})$ are the n-th order contributions and $D_n(\tau)$ are the corresponding growth factors ($D_1 \equiv D$ is the linear growth factor). At any order both the gravitational and velocity potentials can be expressed in terms of these $\boldsymbol{\Psi}^{(n)}$, which are in turn given by the LPT potentials φ_n, e.g.

$$\nabla\cdot\boldsymbol{\Psi}^{(1)} = \nabla^2\varphi_1 = -\delta\,, \tag{7.33}$$

$$\nabla\cdot\boldsymbol{\Psi}^{(2)} = \nabla^2\varphi_2 = -\mathscr{G}_2(\varphi_1)\,. \tag{7.34}$$

Any set of linearly independent operators induced by gravity must therefore be connected to combinations of the LPT potentials. In order to guarantee that these are still

Galilean invariant, we can generalise the definition of the Galileons to (Chan et al., 2012)

$$\mathscr{G}_2(A,B) \equiv \nabla_{ij}A\nabla_{ij}B - \nabla^2 A\nabla^2 B\,, \tag{7.35}$$

and similarly for $\mathscr{G}_3(A,B,C)$. From Eq. (7.33) we have $\varphi_1 = -\Phi_L$, so that the first new combination appears at third order of perturbation theory: $\mathscr{G}_2(\varphi_2,\varphi_1)$. Evaluating this Galileon using Eq. (7.34) shows that it is precisely related to the only gravity induced operator that we previously identified at third-order, i.e. $\Delta_3\mathscr{G}_2 = -4/7\,\mathscr{G}_2(\varphi_2,\varphi_1)$.

Following this line of argument we can now easily determine the additional operators at fourth order: apart from the combination $\delta\mathscr{G}_2(\varphi_2,\varphi_1)$, we can construct the following three invariants out of the LPT potentials

$$\mathscr{G}_2(\varphi_2,\varphi_2)\,,\ \mathscr{G}_2(\varphi_3,\varphi_1)\,,\ \mathscr{G}_3(\varphi_2,\varphi_1,\varphi_1)\,. \tag{7.36}$$

However, beyond second order the LPT solutions are no longer purely potential anymore and at third order in particular it consists of two scalar and a vector potential, all with different time dependencies:

$$\boldsymbol{\Psi}(\boldsymbol{q},\tau)\big|_{\rm 3rd} = D_3^{(a)}(\tau)\,\nabla\varphi_3^{(a)} + D_3^{(b)}(\tau)\,\nabla\varphi_3^{(b)} + D_3^{(c)}(\tau)\,\nabla\times\boldsymbol{A}_3\,, \tag{7.37}$$

where in an Einstein-de Sitter (EdS) universe the growth factors are given by $D_3^{(a)}(\tau) = 1/18\,D(\tau)^3$, $D_3^{(b)}(\tau) = 5/42\,D(\tau)^3$ and $D_3^{(c)}(\tau) = 1/14\,D(\tau)^3$. The potentials, on the other hand, satisfy the following relations (Buchert, 1994)

$$\nabla^2\varphi_3^{(a)} = -\mathscr{G}_3(\varphi_1)\,, \tag{7.38}$$

$$\nabla^2\varphi_3^{(b)} = -\mathscr{G}_2(\varphi_2,\varphi_1)\,, \tag{7.39}$$

$$[\nabla\times(\nabla\times\boldsymbol{A}_3)]_i = -\epsilon_{ijk}\left(\nabla_{jn}\varphi_2\right)\left(\nabla_{kn}\varphi_1\right), \tag{7.40}$$

where ϵ_{ijk} denotes the fully anti-symmetric Levi-Civita symbol. The combination of the third and first order LPT potentials is thus made up of three pieces and factoring out $D(\tau)^3$ from the EdS solutions, we define

$$\mathscr{G}_2(\varphi_3,\varphi_1) \equiv \frac{1}{18}\mathscr{G}_2(\varphi_3^{(a)},\varphi_1) + \frac{5}{42}\mathscr{G}_2(\varphi_3^{(b)},\varphi_1) + \frac{1}{14}\nabla_i\left(\nabla\times\boldsymbol{A}_3\right)_j\nabla_{ij}\varphi_1\,. \tag{7.41}$$

Due to the different time dependencies we should in principle allow these three pieces to enter the bias basis individually, but in practice the departures from the EdS growth factors are small ($\lesssim 0.1$), so for all purposes of this chapter we are safe to ignore this complication.

To conclude, our choice of a complete Galilean invariant basis for the evolved galaxy perturbations is given by a set of 15 operators up to fourth order, which are summarised

Table 7.1 Overview of basis operators for local and non-local galaxy bias, along with their associated bias parameters. Each single column presents a different order in SPT and we have categorised operators into groups, which are local and non-local in the linear gravitational potential (left and middle columns, respectively), and which contain higher than second derivatives of the potentials. Note that for reasons of consistency, we have relabeled some bias parameters compared to Chan et al. (2012) — what was formerly $\gamma_3^\times$ is now $\gamma_2^\times$ and γ_3^- has become $-7/4\,\gamma_{21}$.

Lagrangian				(Additional) Eulerian		Higher-derivative	
1st	2nd	3rd	4th	3rd	4th	3rd	4th
$b_1\,\delta$	$b_2\,\delta^2$	$b_3\,\delta^3$	$b_4\,\delta^4$	$\gamma_{21}\,\mathcal{G}_2(\varphi_2,\varphi_1)$	$\gamma_{21}^\times\,\delta\,\mathcal{G}_2(\varphi_2,\varphi_1)$	$\beta_\delta^{(1)}\,\nabla^2\delta$	$\beta_{\delta^2}^{(1,1)}\,\nabla^2\delta^2$
	$\gamma_2\,\mathcal{G}_2(\Phi_v)$	$\gamma_2^\times\,\delta\,\mathcal{G}_2(\Phi_v)$	$\gamma_2^{\times\times}\,\delta^2\,\mathcal{G}_2(\Phi_v)$		$\gamma_{211}\,\mathcal{G}_3(\varphi_2,\varphi_1,\varphi_1)$		$\beta_{\delta^2}^{(1,2)}\,(\nabla\delta)^2$
		$\gamma_3\,\mathcal{G}_3(\Phi_v)$	$\gamma_3^\times\,\delta\,\mathcal{G}_3(\Phi_v)$		$\gamma_{22}\,\mathcal{G}_2(\varphi_2,\varphi_2)$		$\beta_{\mathcal{G}_2}^{(1,1)}\,\nabla^2\mathcal{G}_2(\Phi_v)$
			$\gamma_2^{\mathrm{sq}}\,\mathcal{G}_2(\Phi_v)^2$		$\gamma_{31}\,\mathcal{G}_2(\varphi_3,\varphi_1)$		$\beta_{\mathcal{G}_2}^{(1,2)}\,\mathcal{G}_2(\nabla\Phi_v)$

in the first two columns of Table 7.1. We have separated the Lagrangian operators (first column), which are local in $\nabla_{ij}\Phi_L$, from those that contain non-linear corrections to the gravitational and velocity potentials (middle column). With respect to the number of basis operators we are thus in agreement with Desjacques et al. (2018) (whereas two operators are missing in Mirbabayi et al. (2015) and one operator in Assassi et al. (2014)) and in fact they can be shown to be equivalent via the relations given in App. 7.A.2.

7.2.4 Higher-derivative galaxy bias

Although some of the basis operators we derived in Sec. 7.2.3 are non-local in the matter fluctuations and gravitational potential, we made the central assumption that the formation of galaxies depends only on the value of these operators at a single point in space. On small scales this approximation must break down because galaxies form due to matter collapsing from a finite region of size $\sim R_*$, which is of the order of the Lagrangian radius of the host dark matter halo. We will show that this effect in general leads to a running of the local bias parameters, which induces a scale dependence that goes as powers of $\sim (k\,R_*)^2$.

Let us begin with a linearly biased galaxy density field. In order to account for an environment dependence we sum up all contributions around the point $\boldsymbol{x}$, modulated by some unknown kernel function F_δ (Desjacques et al., 2018). In essence, we are making the following replacement:

$$b_1\,\delta(\boldsymbol{x}) \longrightarrow \int \mathrm{d}^3y\, F_\delta(y)\,\delta(\boldsymbol{x}+\boldsymbol{y})\,, \tag{7.42}$$

where we require that $F_\delta(y)$ goes to zero if $y \gg R_*$, so that we recover the local bias relation on large scales. Statistical homogeneity further demands that F_δ cannot depend on $\boldsymbol{x}$ itself, while the absence of any preferred directions implies a dependence on the modulus of $\boldsymbol{y}$ only. A Taylor expansion around $\boldsymbol{x}$ gives

$$\delta(\boldsymbol{x}+\boldsymbol{y}) = \delta(\boldsymbol{x}) + \boldsymbol{y}\cdot\nabla\delta(\boldsymbol{x}) + \frac{1}{2}\left(\boldsymbol{y}\cdot\nabla\right)^2\delta(\boldsymbol{x}) + \dots\,, \tag{7.43}$$

and when plugged into Eq. (7.42) we obtain

$$\int \mathrm{d}^3y\, F_\delta(y)\,\delta(\boldsymbol{x}+\boldsymbol{y}) = \left[\int \mathrm{d}^3y\, F_\delta(y)\right]\delta(\boldsymbol{x}) + \left[\frac{1}{6}\int \mathrm{d}^3y\, F_\delta(y)\,y^2\right]\nabla^2\delta(\boldsymbol{x}) + \dots\,, \tag{7.44}$$

where certain contractions between $\boldsymbol{y}$ and the derivative operator vanish after integration (this can be most easily seen when transforming to Fourier space and performing the angular integrations over $\boldsymbol{y}$). In particular, the linear term $\boldsymbol{y}\cdot\nabla\delta$ is explicitly forbidden by statistical isotropy. As the kernel function is unknown we can define its moments in the square brackets as new, free parameters. The first term is simply the linear bias

parameter b_1, whereas the second one has dimension of length squared and which we thus define as $b_1\,\beta_\delta^{(1)}\,R_*^2$, such that

$$\int \mathrm{d}^3y\, F_\delta(y)\,\delta(\boldsymbol{x}+\boldsymbol{y}) = b_1\left[\delta(\boldsymbol{x}) + \beta_\delta^{(1)}\,R_*^2\,\nabla^2\delta(\boldsymbol{x}) + \ldots\right]. \tag{7.45}$$

We now see that the non-locality of galaxy formation introduces new operators, which involve higher than second derivatives of the gravitational potential. Any uneven number of derivatives requires a contraction, which would produce a preferred direction, so Eq. (7.45) leads to the following scale dependent linear bias parameter in Fourier space:

$$b_1(k) = b_1\left[1 + \sum_{n=1}(-1)^n\beta_\delta^{(n)}\left(\frac{k}{k_*}\right)^{2n}\right], \tag{7.46}$$

with $k_* \equiv 1/R_*$. On large scales $k \ll k_*$ the non-local contributions are clearly suppressed and we count each additional derivative acting on the gravitational or velocity potentials as an increase of the SPT order by one. The second term in Eq. (7.45) would thus be considered as third order.

Which additional operators enter at even higher orders of SPT? All of the basis elements at second order and above in Table 7.1 are composites of lower-order operators, and for a generic combination $\mathscr{O} = A_{(1)}\cdots A_{(n)}$ (which can include contractions between the $A_{(n)}$ as for instance $\nabla_{ij}\Phi\,\nabla_{ij}\Phi$) we can generalise Eq. (7.42) as follows (Desjacques et al., 2018):

$$b_{\mathscr{O}}\,\mathscr{O}(\boldsymbol{x}) \longrightarrow \int \mathrm{d}^3y_1\cdots\mathrm{d}^3y_n\,F_{\mathscr{O}}(\boldsymbol{y}_1,\ldots,\boldsymbol{y}_n)\prod_{i=1}^{n}A_{(i)}(\boldsymbol{x}+\boldsymbol{y}_i), \tag{7.47}$$

where $F_{\mathscr{O}}$ can depend only on the moduli and relative orientations of the vectors $\boldsymbol{y}_1,\ldots,\boldsymbol{y}_n$. For the particular case of δ^2 we get after Taylor expanding

$$b_2\,\delta^2(\boldsymbol{x}) \longrightarrow b_2\left[\delta^2(\boldsymbol{x}) + \beta_{\delta^2}^{(1,1)}\,R_*^2\left(\nabla\delta(\boldsymbol{x})\right)^2 + \beta_{\delta^2}^{(1,2)}\,R_*^2\,\nabla^2\delta^2(\boldsymbol{x}) + \ldots\right], \tag{7.48}$$

where we have identified

$$b_2 \equiv \int \mathrm{d}^3y_1\mathrm{d}^3y_2\,F_{\delta^2}(y_1,y_2,\mu_{12}), \tag{7.49}$$

$$b_2\,\beta_{\delta^2}^{(1,1)}\,R_*^2 \equiv -\frac{1}{6}\int \mathrm{d}^3y_1\mathrm{d}^3y_2\,F_{\delta^2}(y_1,y_2,\mu_{12})\,|\boldsymbol{y}_1-\boldsymbol{y}_2|^2, \tag{7.50}$$

$$b_2\,\beta_{\delta^2}^{(1,2)}\,R_*^2 \equiv \frac{1}{12}\int \mathrm{d}^3y_1\mathrm{d}^3y_2\,F_{\delta^2}(y_1,y_2,\mu_{12})\left(y_1^2+y_2^2\right). \tag{7.51}$$

As above this can be shown by transforming into Fourier space, and furthermore we have used that $\delta\,\nabla^2\delta = 1/2\,\nabla^2\delta^2 - (\nabla\delta)^2$. In a similar manner we can derive the higher-derivative corrections for $\mathscr{G}_2$, which gives another two operators up to fourth order. In

total, we should then extend our local bias basis by the following five operators

$$\begin{aligned} &\text{3rd:} \quad \nabla^2\delta\,, \\ &\text{4th:} \quad \nabla^2\delta^2\,,\ \left(\nabla\delta\right)^2\,,\ \nabla^2\mathscr{G}_2(\Phi_v)\,,\ \mathscr{G}_2(\nabla_i\Phi_v,\nabla_i\Phi_v)\,, \end{aligned} \tag{7.52}$$

which makes five less than the set given in Desjacques et al. (2018).

7.2.5 Local Lagrangian bias relations

In Sec. 7.2.3 we argued that all bias operators derive from second derivatives of the gravitational potential and its time evolution. Consequently, these operators are generally non-local in the matter density. Moreover, this implies that even if the galaxy abundance on some initial (Lagrangian) time slice is truly locally related to matter, i.e. $\delta_g = \sum \bar{b}_n \delta^n$, gravitational evolution will inevitably invalidate this assumption at any later point in time.

However, Chan et al. (2012) showed that in this case it is possible to express all parameters, which are non-local in δ, in terms of the local ones. Considering only the fastest growing modes, they find:

$$\gamma_2 = -\frac{2}{7}\left(b_1 - 1\right)\,, \quad \gamma_3 \ = \frac{11}{63}\left(b_1 - 1\right)\,, \tag{7.53}$$

$$\gamma_{21} = \frac{33}{126}\left(b_1 - 1\right)\,, \quad \gamma_2^{\times} = -\frac{2}{7}\,b_2\,, \tag{7.54}$$

which we will denote as the *local Lagrangian bias* relations. Following the procedure outlined in Chan et al. (2012) it is possible to derive similar relations for the non-local fourth-order parameters. While this would go beyond the scope of the current chapter, we note that the corresponding terms are implicitly generated by the recursive evolution equation presented in Sec. 7.3.3.

7.3 The multi-point propagator formalism

We now return to the main idea presented in Sec. 7.2.2: in order to guarantee that the bias parameters corresponding to our Galilean basis operators are observable quantities, we should construct the galaxy density field out of multi-point propagators,

$$\left\langle \frac{\partial^n \delta_g(\boldsymbol{k})}{\partial\delta_L(\boldsymbol{k}_1)\cdots\partial\delta_L(\boldsymbol{k}_n)} \right\rangle\,, \tag{7.55}$$

as opposed to the usual kernel functions that we obtain from taking derivatives at $\delta_L = 0$, i.e. $\left[\partial^n\delta_g(\boldsymbol{k})/\partial\delta_L(\boldsymbol{k}_1)\cdots\partial\delta_L(\boldsymbol{k}_n)\right]_0$. Because of the Gaussianity of the initial matter fluctuations δ_L, this is equivalent with expanding the galaxy density in terms of

Wiener-Hermite functionals, which was first considered by Szalay (1988) and further developed in Matsubara (1995). We will begin by formalising this kind of expansion, which will be followed up by the computation of the galaxy propagators at the level of the initial conditions, as well as a derivation of their time evolution.

7.3.1 Generalised Wiener-Hermite expansion

For a linear Gaussian dark matter density field, the probability density function (PDF) for a mode in Fourier space is given by

$$\mathscr{P}[\delta_L] = N \exp\left[-\frac{1}{2}\int_{\boldsymbol{q}} \frac{|\delta_L(\boldsymbol{q})|^2}{P_L(q)}\right], \tag{7.56}$$

with normalisation factor N. The n-th generalised Wiener-Hermite functional $\mathscr{H}_n$ is then defined by taking n functional derivatives of the PDF (Matsubara, 1995):

$$\mathscr{H}_n(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n) \equiv \frac{(-1)^n}{\mathscr{P}[\delta_L]} \frac{\partial^n \mathscr{P}[\delta_L]}{\partial\delta_L(\boldsymbol{k}_1)\cdots\partial\delta_L(\boldsymbol{k}_n)} \prod_{i=1}^{n} P_L(\boldsymbol{k}_i)\,. \tag{7.57}$$

Using this definition, we obtain the following first three functionals (suppressing the momentum arguments):

$$\begin{aligned}
\mathscr{H}_1 &= \delta_L(-\boldsymbol{k})\,,\\
\mathscr{H}_2 &= \delta_L(-\boldsymbol{k}_1)\,\delta_L(-\boldsymbol{k}_2) - \langle\delta_L(\boldsymbol{k}_1)\,\delta_L(\boldsymbol{k}_2)\rangle\,,\\
\mathscr{H}_3 &= \delta_L(-\boldsymbol{k}_1)\,\delta_L(-\boldsymbol{k}_2)\,\delta_L(-\boldsymbol{k}_3) - \Big[\langle\delta_L(\boldsymbol{k}_1)\,\delta_L(\boldsymbol{k}_2)\rangle\,\delta_L(-\boldsymbol{k}_3) + \text{cyc.}\Big]\,,
\end{aligned} \tag{7.58}$$

and like the original Hermite polynomials they satisfy an orthogonality relation, which can be shown to be (see Matsubara, 1995, and App. 7.B.1):

$$\begin{aligned}
\langle\mathscr{H}_n(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n)\,\mathscr{H}_m^*(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n)\rangle = (2\pi)^{3n}\,\delta_{nm} \prod_{i=1}^{n} P_L(\boldsymbol{k}_i)\\
\times\Big[\delta_D(\boldsymbol{k}_1-\boldsymbol{q}_1)\cdots\delta_D(\boldsymbol{k}_n-\boldsymbol{q}_n) + \text{sym.}\Big]
\end{aligned} \tag{7.59}$$

where $\mathscr{H}_m^*$ denotes the complex conjugate and 'sym.' stands for the remaining $n!-1$ combinations of the arguments.

In Fourier space the contribution at each order is written as a convolution over $\mathscr{H}_n$, such that

$$\delta_g(\boldsymbol{k}) = \sum_n \frac{(2\pi)^3}{n!} \int_{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n} \delta_D(\boldsymbol{k}-\boldsymbol{k}_{1\cdots n})\,\Gamma_g^{(n)}(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n)\,\mathscr{H}_n^*(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n)\,, \tag{7.60}$$

where the kernel functions $\Gamma_g^{(n)}(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n)$ can be interpreted as the corresponding bias parameters, which now have acquired a scale dependence. Multiplying both sides of Eq. (7.60) with $\mathscr{H}_m$ and using the orthogonality relation (Eq. 7.59), we see that

$$\langle \mathscr{H}_n(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n)\,\delta_g(\boldsymbol{k})\rangle = (2\pi)^3\,\Gamma_g^{(n)}(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n)\,\delta_D(\boldsymbol{k}-\boldsymbol{k}_{1\cdots n})\prod_{i=1}^{n} P_L(\boldsymbol{k}_i)\,. \tag{7.61}$$

On the other hand, we can derive a different relation by plugging in Eq. (7.57) and replacing the ensemble average by its definition — the functional integral of all modes δ_L over their joint PDF:

$$\begin{aligned}\frac{\langle \mathscr{H}_n(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n)\,\delta_g(\boldsymbol{k})\rangle}{\prod_{i=1}^{n} P_L(\boldsymbol{k}_i)} &= (-1)^n \int \mathscr{D}[\delta_L]\left[\frac{\partial^n \mathscr{P}[\delta_L]}{\partial\delta_L(\boldsymbol{k}_1)\cdots\partial\delta_L(\boldsymbol{k}_n)}\right]\delta_g(\boldsymbol{k}) \\ &= \int \mathscr{D}[\delta_L]\,\mathscr{P}[\delta_L]\,\frac{\partial^n \delta_g(\boldsymbol{k})}{\partial\delta_L(\boldsymbol{k}_1)\cdots\partial\delta_L(\boldsymbol{k}_n)} \\ &= \left\langle \frac{\partial^n \delta_g(\boldsymbol{k})}{\partial\delta_L(\boldsymbol{k}_1)\cdots\partial\delta_L(\boldsymbol{k}_n)}\right\rangle,\end{aligned} \tag{7.62}$$

where we have integrated by parts n times in going from the first to the second line. Thus, we have shown that the kernels of the Wiener-Hermite expansion are indeed multi-point propagators:

$$\left\langle \frac{\partial^n \delta_g(\boldsymbol{k})}{\partial\delta_L(\boldsymbol{k}_1)\cdots\partial\delta_L(\boldsymbol{k}_n)}\right\rangle = (2\pi)^3\,\Gamma_g^{(n)}(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n)\,\delta_D(\boldsymbol{k}-\boldsymbol{k}_{1\cdots n})\,, \tag{7.63}$$

and as argued above this ensures that the expansion does not contain any diverging or cutoff dependent contributions. In particular, we can expect that in the large-scale limit the first order Γ_g is given by a single constant parameter — the linear bias parameter b_1 — as shown by observations.

We also note that Eq. (7.63) is equivalent with the definition of multi-point propagators in Bernardeau et al. (2008), which, however, considered the non-linear matter fluctuations in place of the galaxy fluctuations. They showed that the multi-point propagators function as basic building blocks for constructing arbitrary N-point spectra. The same should therefore hold for the $\Gamma_g^{(n)}$ and we will use this fact to compute the galaxy power spectrum and bispectrum in Sec. 7.4.

7.3.2 Galaxy propagators

Having seen that the galaxy propagators correspond to the observed bias parameters, our task now will be to determine how they vary for different wave number configurations. This scale dependence cannot be arbitrary and is in fact determined by the relation between the galaxy over-density and the gravitational and velocity potential of the underlying dark matter field. Hence, as a starting point we will use the bias expan-

sion presented in Sec. 7.2.3 and 7.2.4 and apply Eq. (7.63) in order to compute the galaxy propagators. Additionally, we will work in Lagrangian space, so that non-linearities in the dark matter over-density can be ignored, which corresponds to setting all SPT kernels F_n and G_n for $n \geq 2$ to zero.

The one-point propagator

As a first step let us consider a single derivative of the linear matter over-density, which we can rewrite using Eq. (7.62):

$$\left\langle \frac{\partial \delta_L(\boldsymbol{k})}{\partial \delta_L(\boldsymbol{k}')} \right\rangle = \frac{\langle \mathcal{H}_1(\boldsymbol{k}')\, \delta_L(\boldsymbol{k}) \rangle}{P_L(k)} = \frac{\langle \delta_L(-\boldsymbol{k}')\, \delta_L(\boldsymbol{k}) \rangle}{P_L(k)} = (2\pi)^3\, \delta_D(\boldsymbol{k} - \boldsymbol{k}')\,, \tag{7.64}$$

where we have plugged in the first Wiener-Hermite functional from Eq. (7.58). Taking instead a derivative of two powers of δ_L would yield a result proportional to $\langle \delta_L \rangle$, which evaluates to zero. More generally, as each derivative cancels exactly one factor of δ_L, we see that odd (even) numbered propagators can only contain terms stemming from odd (even) orders of the bias expansion.

Thus, to complete the computation of $\Gamma_g^{(1)}$, we determine the derivative of a generic third-order term $\mathcal{O}_B^{(3)}$, which can be written as the convolution

$$\mathcal{O}_B^{(3)}(\boldsymbol{k}) = (2\pi)^3 \int\limits_{\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3} \delta_D(\boldsymbol{k} - \boldsymbol{k}_{123})\, \mathcal{K}_B^{(3)}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)\, \delta_L(\boldsymbol{k}_1)\, \delta_L(\boldsymbol{k}_2)\, \delta_L(\boldsymbol{k}_3)\,, \tag{7.65}$$

with $B \in \{\delta^3, \delta\mathcal{G}_2, \mathcal{G}_3, \mathcal{G}_2(\varphi_2, \varphi_1)\}$ and the kernels $\mathcal{K}_B^{(3)}$ are given in Eqs. (7.117) to (7.120). Making use of the symmetry of $\mathcal{K}_B^{(3)}$, we obtain:

$$\left\langle \frac{\partial \mathcal{O}_B^{(3)}(\boldsymbol{k})}{\partial \delta_L(\boldsymbol{k}')} \right\rangle' = 3 \int_{\boldsymbol{q}} \mathcal{K}_B^{(3)}(\boldsymbol{k}, \boldsymbol{q}, -\boldsymbol{q})\, P_L(q)\,, \tag{7.66}$$

where a prime on the angle brackets indicates that we have dropped the factor of $(2\pi)^3$ as well as the Dirac delta distribution, which commonly appear when taking ensemble averages, from the right-hand side of the equation. Therefore, it represents the direct contribution to $\Gamma_g^{(1)}$ due to third-order bias terms. As $L(\boldsymbol{k}, \boldsymbol{q}, -\boldsymbol{q}) = 0$, the only nontrivial cases we need to consider are for $B = \delta\mathcal{G}_2, \mathcal{G}_2(\varphi_2, \varphi_1)$, and plugging in the kernel for the former, we notice that its contribution, like that for δ^3, is actually scale independent:

$$\begin{aligned} \left\langle \frac{\partial \mathcal{O}_{\delta\mathcal{G}_2}^{(3)}(\boldsymbol{k})}{\partial \delta_L(\boldsymbol{k}')} \right\rangle' &= 2 \int_{\boldsymbol{q}} K(\boldsymbol{k}, \boldsymbol{q})\, P_L(q) = \frac{1}{2\pi^2} \int q^2 \mathrm{d}q \int_{-1}^{1} \mathrm{d}\mu \left(\mu^2 - 1\right) P_L(q) \\ &= -\frac{4}{3} \int \mathrm{d}q \frac{q^2\, P_L(q)}{2\pi^2} = -\frac{4}{3}\sigma^2\,. \end{aligned} \tag{7.67}$$

For $B = \mathcal{G}_2(\varphi_2, \varphi_1)$ on the other hand, we get

$$\left\langle \frac{\partial \mathcal{O}^{(3)}_{\mathcal{G}_2(\varphi_2,\varphi_1)}(\boldsymbol{k})}{\partial \delta_L(\boldsymbol{k}')} \right\rangle' = 2 \int_{\boldsymbol{q}} K(\boldsymbol{k}-\boldsymbol{q},\boldsymbol{q})\, K(\boldsymbol{k},\boldsymbol{q})\, P_L(q)\,, \tag{7.68}$$

which can be shown to vanish in the large-scale limit as $(k/q)^2$. Collecting all previous results, the first galaxy propagator is given by

$$\Gamma^{(1)}_g(\boldsymbol{k}) = \bar{b}_1 + \left[\frac{1}{2}\bar{b}_3 - \frac{4}{3}\bar{\gamma}^{\times}_2\right]\sigma^2 + 2\,\bar{\gamma}_{21} \int_{\boldsymbol{q}} K(\boldsymbol{k}-\boldsymbol{q},\boldsymbol{q})\, K(\boldsymbol{k},\boldsymbol{q})\, P_L(q)\,, \tag{7.69}$$

demonstrating that up to fourth order in the bias expansion, we only require two parameters to describe the scale dependence of $\Gamma^{(1)}_g$. Accordingly, we make the following two redefinitions:

$$b_1 \equiv \bar{b}_1 + \left[\frac{1}{2}\bar{b}_3 - \frac{4}{3}\bar{\gamma}^{\times}_2\right]\sigma^2\,, \tag{7.70}$$

$$\gamma_{21} \equiv \bar{\gamma}_{21}\,, \tag{7.71}$$

and the behaviour of the γ_{21} integral in the limit $k \to 0$, indeed ensures that $\Gamma^{(1)}_g$ corresponds to the linear bias parameter on large scales. As explained in Sec. 7.2.2, the redefinitions above will keep changing when higher orders from the bias expansion are included. In addition, further terms stemming from the non-local relation between the galaxy and matter densities will appear, which are, however, suppressed by higher powers of k. Finally, considering higher-derivative bias, we note that only the third-order term in Eq. (7.52) can contribute to the one-point propagator. This gives rise to the correction,

$$\Gamma^{(1)}_{g,\nabla}(\boldsymbol{k}) = -b_1\,\beta^{(1)}_{\delta}\left(\frac{k}{k_*}\right)^2\,, \tag{7.72}$$

and introduces a scale dependence of the linear bias parameter, as already shown in Sec. 7.2.4.

The two-point propagator

In a similar manner we can now derive all remaining multi-point propagators. The two-point propagator gets contributions from second and fourth order bias operators, which, analogous to Eq. (7.65), we write as the symmetric kernels $\mathcal{K}^{(2)}_B$ and $\mathcal{K}^{(4)}_B$. The corresponding expressions are given in Eqs. (7.114) and (7.115), and Eqs. (7.124) to (7.131), respectively. Differentiating a generic second or fourth order contribution twice results in

$$\left\langle \frac{\partial^2 \mathcal{O}^{(2)}_B(\boldsymbol{k})}{\partial \delta_L(\boldsymbol{k}_1)\,\partial \delta_L(\boldsymbol{k}_2)} \right\rangle' = 2\,\mathcal{K}^{(2)}_B(\boldsymbol{k}_1,\boldsymbol{k}_2)\,, \tag{7.73}$$

and

$$\left\langle \frac{\partial^2 \mathcal{O}_B^{(4)}(\boldsymbol{k})}{\partial\delta_L(\boldsymbol{k}_1)\partial\delta_L(\boldsymbol{k}_2)} \right\rangle' = 12\int_{\boldsymbol{q}} \mathcal{K}_B^{(4)}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{q},-\boldsymbol{q})\,P_L(q)\,, \tag{7.74}$$

such that after plugging in the kernels we recognise that the loop corrections to the two-point galaxy propagator consist only of the additional Eulerian basis operators that stem from non-linear evolution of the potentials:

$$\begin{aligned}\Gamma_g^{(2)}(\boldsymbol{k}_1,\boldsymbol{k}_2) = b_2 + 2\gamma_2\,K(\boldsymbol{k}_1,\boldsymbol{k}_2) + 12\int_{\boldsymbol{q}}\Big[&\gamma_{21}^{\times}\,\mathcal{K}^{(4)}_{\delta\mathcal{G}_2(\varphi_2,\varphi_1)} + \gamma_{211}\,\mathcal{K}^{(4)}_{\mathcal{G}_3(\varphi_2,\varphi_1,\varphi_1)} \\ &+ \gamma_{22}\,\mathcal{K}^{(4)}_{\mathcal{G}_2(\varphi_2,\varphi_2)} + \gamma_{31}\,\mathcal{K}^{(4)}_{\mathcal{G}_2(\varphi_3,\varphi_1)}\Big]\,P_L(q)\,.\end{aligned} \tag{7.75}$$

Explicit expressions for the kernels, whose momentum dependence we have suppressed in the integral above, are given in App. 7.A.1. All other bias parameters disappear after we have identified the following two renormalised combinations by grouping together terms with the same scale dependence,

$$b_2 \equiv \bar{b}_2 + \left[\frac{1}{2}\bar{b}_4 - \frac{16}{3}\bar{\gamma}_2^{\times\times} + \frac{32}{15}\bar{\gamma}_{21}^{\times} + \frac{64}{15}\bar{\gamma}_2^{\mathrm{sq}}\right]\sigma^2\,, \tag{7.76}$$

$$\gamma_2 \equiv \bar{\gamma}_2 + \left[\bar{\gamma}_2^{\times\times} + \frac{2}{5}\bar{\gamma}_{21}^{\times} - \frac{1}{2}\bar{\gamma}_3^{\times} + \frac{8}{15}\bar{\gamma}_2^{\mathrm{sq}}\right]\sigma^2\,. \tag{7.77}$$

These are the observable quadratic and tidal field bias parameters, which match those determined in Assassi et al. (2014) up to the signs of the $\gamma_{21}^{\times}$ and $\gamma_3^{\times}$ terms in the expression for γ_2 (note that in their notation $b_{\Gamma_3\delta} = -4/7\,\gamma_{21}^{\times}$ and $b_{\mathcal{G}_3\delta} = \gamma_3^{\times}$). As already observed for the one-point propagator, the loop integrals in Eq. (7.75) scale as $|\boldsymbol{k}_1+\boldsymbol{k}_2|^2/q^2$ in the large-scale limit and are therefore suppressed compared to the first two (tree-level) terms. The higher-derivative corrections have a similar scale dependence and lead to the following four additional terms to be included in Eq. (7.75):

$$\Gamma^{(2)}_{g,\nabla}(\boldsymbol{k}_1,\boldsymbol{k}_2) = -b_2\left[\beta^{(1,1)}_{\delta^2}\frac{k_{12}^2}{k_*^2} + \beta^{(1,2)}_{\delta^2}\frac{\boldsymbol{k}_1\cdot\boldsymbol{k}_2}{k_*^2}\right] - 2\gamma_2\left[\beta^{(1,1)}_{\mathcal{G}_2}\frac{k_{12}^2}{k_*^2} + \beta^{(1,2)}_{\mathcal{G}_2}\frac{\boldsymbol{k}_1\cdot\boldsymbol{k}_2}{k_*^2}\right]K(\boldsymbol{k}_1,\boldsymbol{k}_2)\,. \tag{7.78}$$

The expressions in the square brackets can be interpreted as a scale dependence of the quadratic and tidal field bias parameters.

The three-point propagator

Finally, for the three-point propagator, we compute three derivatives of $\mathcal{O}_B^{(3)}$, giving

$$\left\langle \frac{\partial^3 \mathcal{O}_B^{(3)}(\boldsymbol{k})}{\partial\delta_L(\boldsymbol{k}_1)\partial\delta_L(\boldsymbol{k}_2)\partial\delta_L(\boldsymbol{k}_3)} \right\rangle \simeq 6\,\mathcal{K}_B^{(3)}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)\,. \tag{7.79}$$

At the order of PT we are working in, there are thus no loop integrals and we simply get:

$$\Gamma_g^{(3)}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) = b_3 + 2\gamma_2^{\times}\left[K(\boldsymbol{k}_1,\boldsymbol{k}_2)+\text{cyc.}\right] + 2\gamma_{21}\left[K(\boldsymbol{k}_1,\boldsymbol{k}_2)\,K(\boldsymbol{k}_{12},\boldsymbol{k}_3)+\text{cyc.}\right] + 6\gamma_3\,L(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)\,. \tag{7.80}$$

According to our counting any contribution from higher-derivative bias to the three-point propagator would be of fifth order or higher. For that reason no further corrections have to be included up to the order we are working in.

7.3.3 Time evolution

So far we have considered the bias relation to be imposed on some initial time slice. However, until the time where galaxies are observed, all of the basis operators in Table 7.1 will have evolved and thus developed non-linear corrections that are of higher SPT orders. For instance, while the basis operator $b_1\,\delta$ is first order in the initial conditions, when evaluated at some later time it contributes at all consecutive orders by means of the SPT solutions $\delta^{(n)}$ presented in Eq. (7.5).

Consequently, if we were to compute galaxy multi-point propagators at late times and at a given order in SPT, we would have to account for all non-linear corrections of those operators which are initially of a lower order. As an example let us consider the contribution from $b_2\,\delta^2$ to $\Gamma_g^{(2)}$. Allowing for time evolution means that we get an additional fourth-order contribution $\sim b_2\,\delta^{(2)}\delta^{(2)}$ (amongst others), which leads to the following loop integral:

$$\int_{\boldsymbol{q}} F_2(\boldsymbol{k}_1,\boldsymbol{q})\,F_2(\boldsymbol{k}_2,-\boldsymbol{q})\,P_L(q)\,. \tag{7.81}$$

This expression contains both, terms that are proportional to σ^2 and thus potentially divergent, as well as terms that give a finite contribution to $\Gamma_g^{(2)}$. The former can be absorbed by the renormalisations of b_1, b_2 and γ_2 and must be subtracted from the integral, which, given the number of terms we are dealing with up to fourth order, becomes a very cumbersome procedure.

For that reason we will follow a different approach — we start from the Wiener-Hermite expansion at initial time using the propagators derived in Sec. 7.3.2 and then evolve this expansion instead of resorting to the usual SPT solutions. As we will see below this guarantees that none of the non-linear corrections to the propagators contain divergencies, and thus highlights another advantage of working in the multi-point propagator formalism.

As a first step we write a combined evolution equation for matter, its velocity divergence and galaxies, and for that purpose we define the three-component vector

$$\boldsymbol{\Psi}(\boldsymbol{k},\tau) \equiv \left(\delta(\boldsymbol{k},\tau),\,\theta(\boldsymbol{k},\tau)/f\mathcal{H},\,\delta_g(\boldsymbol{k},\tau)\right). \tag{7.82}$$

In terms of $\mathbf{\Psi}$ and by using the logarithm of the linear growth rate as our new time variable, i.e. $\eta \equiv \ln D(\tau)$, the evolution equations can be recast as (see Chan et al., 2012))

$$\frac{\partial \Psi_a(\boldsymbol{k},\eta)}{\partial \eta} + \Omega_{ab}\,\Psi_b(\boldsymbol{k},\eta) = (2\pi)^3 \int\limits_{\boldsymbol{k}_1,\boldsymbol{k}_2} \delta_D(\boldsymbol{k}-\boldsymbol{k}_{12})\gamma_{abc}(\boldsymbol{k}_1,\boldsymbol{k}_2)\,\Psi_b(\boldsymbol{k}_1,\eta)\,\Psi_c(\boldsymbol{k}_2,\eta)\,, \tag{7.83}$$

where we follow the convention that repeated indices are summed over. The matrix

$$\Omega_{ab} \equiv \frac{1}{2}\begin{bmatrix} 0 & -2 & 0 \\ -3 & 1 & 0 \\ 0 & -2 & 0 \end{bmatrix} \tag{7.84}$$

describes the coupling between densities and velocities, while γ_{abc} encodes the non-linear interactions between different Fourier modes. Its only non-zero components are given by

$$\gamma_{121} = \gamma_{323} = \alpha(\boldsymbol{k}_1,\boldsymbol{k}_2)\,, \tag{7.85}$$

$$\gamma_{222} = \beta(\boldsymbol{k}_1,\boldsymbol{k}_2)\,, \tag{7.86}$$

and $\gamma_{112}(\boldsymbol{k}_1,\boldsymbol{k}_2) = \gamma_{121}(\boldsymbol{k}_2,\boldsymbol{k}_1)$. With Eq. (7.83) we have implicitly assumed that the number of galaxies is conserved troughout all times, i.e. after formation they simply move under the influence of gravity, but do not undergo any mergers (Fry, 1996). In addition, and as already emphasised above, we assume that galaxies and dark matter are comoving, meaning we neglect any kind of velocity bias. For the initial conditions $\phi(\boldsymbol{k}) = \Psi(\boldsymbol{k},\eta=0)$ there thus exists an integral solution to Eq. (7.83), which was derived in Scoccimarro (1998):

$$\begin{aligned}\Psi_a(\boldsymbol{k},\eta) = g_{ab}(\eta)\,\phi(\boldsymbol{k}) + (2\pi)^3 \int_0^\eta \mathrm{d}\eta'\, g_{ab}(\eta-\eta') \\ \times \int\limits_{\boldsymbol{k}_1,\boldsymbol{k}_2} \delta_D(\boldsymbol{k}-\boldsymbol{k}_{12})\gamma_{bcd}(\boldsymbol{k}_1,\boldsymbol{k}_2)\,\Psi_c(\boldsymbol{k}_1,\eta')\,\Psi_d(\boldsymbol{k}_2,\eta')\,.\end{aligned} \tag{7.87}$$

This expression depends on the linear propagator $g_{ab}(\eta)$, which solves the linearised equations of motion (i.e. setting the right-hand side of Eq. (7.83) to zero) and presents a mixture of growing and decaying, as well as time independent modes (Chan et al., 2012; Scoccimarro, 1998):

$$g_{ab}(\eta) = \frac{\mathrm{e}^{\eta}}{5}\begin{bmatrix} 3 & 2 & 0 \\ 3 & 2 & 0 \\ 3 & 2 & 0 \end{bmatrix} - \frac{\mathrm{e}^{-3\eta/2}}{5}\begin{bmatrix} -2 & 2 & 0 \\ 3 & -3 & 0 \\ -2 & 2 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}. \tag{7.88}$$

As we did for the galaxy over-density in Sec. 7.3.1, we now expand $\Psi(\boldsymbol{k},\eta)$ in terms of generalised Wiener-Hermite functionals,

$$\Psi_a(\boldsymbol{k},\eta) = \sum_n \frac{(2\pi)^3}{n!} \int_{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n} \delta_D(\boldsymbol{k}-\boldsymbol{k}_{1\cdots n})\, \Gamma_a^{(n)}(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n;\eta)\, \mathscr{H}_n^*(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n)\,, \tag{7.89}$$

where $\Gamma_3^{(n)} \equiv \Gamma_g^{(n)}$. An equivalent expansion holds at initial time $\eta = 0$ and we denote the corresponding propagators by the symbol $\widehat{\Gamma}$. Since the dark matter and velocity fields are linear at that time, we have $\widehat{\Gamma}_1^{(1)} = 1 = \widehat{\Gamma}_2^{(1)}$, while all higher-order propagators must vanish. The initial galaxy propagators, on the other hand, are given given by the expressions from Sec. 7.3.2. We allow for the possibility that the initial galaxy fluctuations require basis operators up to third order and therefore keep the one-loop correction that appears in Eq. (7.69), whereas we assume that the initial values of the fourth order bias parameters are negligible. However, that does not imply that we ignore their contributions completely because, as discussed in Sec. 7.2.5, time evolution via Eq. 7.83 will generate the one-loop terms in Eq. (7.75) in any case, but with their amplitudes fixed through lower order bias parameters.

By multiplying both sides of Eq. (7.87) with $\mathscr{H}_n$ and taking the ensemble average, we can derive a recursion relation for the time evolved multi-point propagators. Following the steps detailed in App. 7.B, we arrive at the expression:

$$\begin{aligned}\Gamma_a^{(n)}(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n;\eta) &= g_{ab}(\eta)\,\widehat{\Gamma}_b^{(n)}(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n) \\ &\quad + \sum_{m=1}^{n}\sum_{r=r_*}^{m} \frac{1}{r!}\int_0^\eta \mathrm{d}\eta'\, g_{ab}(\eta-\eta')\,\Gamma_b^{(n,m,r)}(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n;\eta')\,,\end{aligned} \tag{7.90}$$

where the quantity $\Gamma_a^{(n,m,r)}$ represents r loop integrals over propagators of orders m and $2r+n-m$:

$$\begin{aligned}\Gamma_a^{(n,m,r)}(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n;\eta') \equiv \int_{\boldsymbol{q}_1,\ldots,\boldsymbol{q}_r} \Big[&\gamma_{abc}(\boldsymbol{k}_{1\cdots m-r}+\boldsymbol{q}_{1\cdots r}, \boldsymbol{k}_{m-r+1\cdots n}-\boldsymbol{q}_{1\cdots r}) \\ &\times \Gamma_c^{(2r+n-m)}(\boldsymbol{k}_{m-r+1},\ldots,\boldsymbol{k}_n,-\boldsymbol{q}_1,\ldots,-\boldsymbol{q}_r;\eta') \\ &\times \Gamma_b^{(m)}(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_{m-r},\boldsymbol{q}_1,\ldots,\boldsymbol{q}_r;\eta') \prod_{i=1}^{r} P_L(\boldsymbol{q}_i) + \text{sym.}(\boldsymbol{k}_i)\Big]\,.\end{aligned} \tag{7.91}$$

The summation over r starts at index

$$r_* = \begin{cases} 0\,, & n > m \\ 1\,, & n = m \\ m-n\,, & n < m \end{cases}\,, \tag{7.92}$$

and sym.($\boldsymbol{k}_i$) stands for the symmetrisation over all possibilities of building a subset of $m-r$ $\boldsymbol{k}$-modes from a total group of n, i.e. $\binom{n}{m-r}$ terms.

Furthermore, Eq. (7.91) illustrates the point we made at the beginning of this section — the time evolved multi-point propagators are free of potentially divergent contributions like those contained in Eq. (7.81), meaning we do not have to perform any additional renormalisation steps. This is a consequence of the scale dependence of the mode coupling kernels γ_{abc}, which contribute to $\Gamma_g^{(n)}$ only via the symmetrised combination $\alpha^s(\boldsymbol{k}_1,\boldsymbol{k}_2) = 1/2[\alpha(\boldsymbol{k}_1,\boldsymbol{k}_2)+\alpha(\boldsymbol{k}_2,\boldsymbol{k}_1)]$. Plugging in the wave vectors from Eq. (7.91) and expanding in inverse powers of $q_{1\cdots r}$, we find that to leading order

$$\alpha^s(\boldsymbol{K}-\boldsymbol{q}_{1\cdots r},\boldsymbol{k}_{1\cdots n}-\boldsymbol{K}-\boldsymbol{q}_{1\cdots r}) \sim \left(\frac{k_{1\cdots n}}{q_{1\cdots r}}\right)^2 + \dots\,, \tag{7.93}$$

where $\boldsymbol{K} = \boldsymbol{k}_{1\cdots m-r}$. In this chapter we are interested in corrections only up to the one-loop level, so we are led to consider expressions of the form

$$\sim k_{1\cdots n}^2 \int_{\boldsymbol{q}} \frac{P_L(q)}{q^2}\,\Gamma^{(m)}(\boldsymbol{k}_1,\dots,\boldsymbol{k}_{m-1},\boldsymbol{q})\,\Gamma^{(2+n-m)}(\boldsymbol{k}_m,\dots,\boldsymbol{k}_n,-\boldsymbol{q})\,, \tag{7.94}$$

where both propagators are evaluated at tree-level. That means the propagators remain finite when the loop momentum q becomes large, which in turn implies that the overall integrand scales, at most, as $1/q^3$ in this limit (using that $P_L(q) \sim 1/q^3$ for $q \to \infty$). This guarantees that any integral of the above type has only a limited sensitivity to the highly non-linear small-scale regime. In particular, no terms involving σ^2 can appear due to time evolution of the multi-point propagators.

For higher than one-loop corrections this argument no longer holds, as each loop adds an additional power spectrum to the integrand, while the scaling with the loop momenta remains the same. However, the scale dependence on the external momenta, i.e. $\sim k_{1\cdots n}^2$, suggests that these terms can be absorbed by redefinitions of the higher-derivative bias parameters, which display the exact same scale dependence, as shown in Sec. 7.2.4. This behaviour of the loop integrals is well known in the context of plain matter perturbations (see Bernardeau et al., 2012), where the small scale sensitivity can be understood in terms of non-zero stress tensor corrections. Indeed, we will show in Sec. 7.4.2 that these are completely degenerate with the contributions from higher-derivative bias.

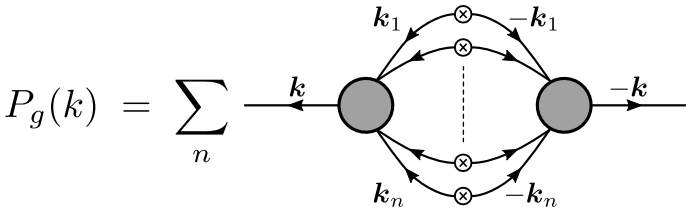


Fig. 7.2 Galaxy power spectrum, reconstructed from multi-point propagators, which are represented by the shaded circles with incoming and outgoing momentum $\boldsymbol{k}$. The sum runs over the number of connected internal lines, each of which produces a linear power spectrum depicted by a crossed circle.

7.4 Power Spectrum and Bispectrum

7.4.1 Reconstructing correlators from multi-point propagators

The multi-point propagators serve as the basic building blocks for computing N-point spectra. This follows easily from the orthogonality relations of the generalised Wiener-Hermite functionals (see App. 7.B.1), and was already shown for the dark matter density in (Bernardeau et al., 2008), whose results we can directly apply to the present case of galaxy clustering.

In particular, by evaluating $\langle \Psi_a(\boldsymbol{k}) \Psi_b(\boldsymbol{k}') \rangle$ and using Eq. (7.59) one finds that the power spectrum is given by a series of two contracted multi-point propagators of the same order. Diagrammatically this can be represented by gluing together two of the objects shown in Fig. 7.1, where each combination of the incoming lines gives rise to a (linear) power spectrum. This is demonstrated in Fig. 7.2 and as the shaded circles include all vertex loop corrections (i.e. vertex renormalisations) we only need to consider one distinct diagram for the power spectrum at one-loop level, compared to the usual two in the standard treatment. The galaxy power spectrum is thus given by Bernardeau et al. (2008)

$$P_g(k) = \left[\Gamma_g^{(1)}(k)\right]^2 P_L(k) + \frac{1}{2}\int_{\boldsymbol{q}} \left[\Gamma_g^{(2)}(\boldsymbol{k}-\boldsymbol{q},\boldsymbol{q})\right]^2 P_L(|\boldsymbol{k}-\boldsymbol{q}|)\, P_L(q)\,, \tag{7.95}$$

where $\Gamma_g^{(1)}$ is evaluated up to one-loop order, while tree-level terms are sufficient for $\Gamma_g^{(2)}$.

We proceed in a similar manner for the bispectrum, which is obtained from the three-point correlator $\langle \Psi_a(\boldsymbol{k}_1) \Psi_b(\boldsymbol{k}_2) \Psi_c(\boldsymbol{k}_3) \rangle$ and application of Eq. (7.152). The complete solution can be found in Bernardeau et al. (2008), but its diagrammatic depiction in Fig. 7.3 is straightforward — a combination of three multi-point propagators with a varying number of connecting lines between each pair (maximally one such pair is allowed to be disconnected to avoid an overall disconnected graph). Taking care of the appropriate symmetry factors that arise in these various combinations, we arrive at the

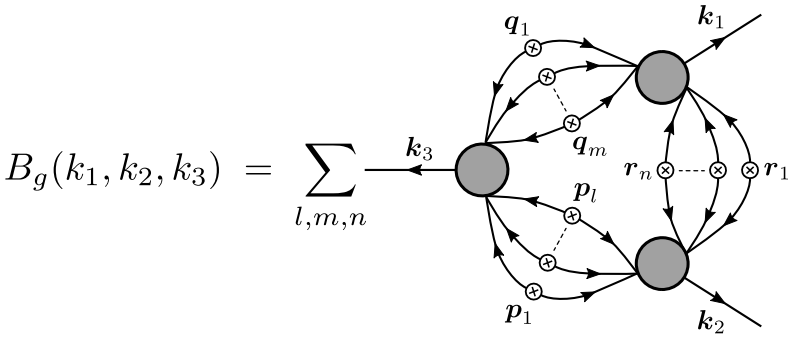


Fig. 7.3 Galaxy bispectrum, expressed through multi-point propagators. The sum runs over the number of connected internal lines of each pair of propagators (shaded circles). At most one of the three indices can be zero, so that the overall diagram remains a connected graph.

following result (Bernardeau et al., 2008):

$$\begin{aligned} B_g(k_1,k_2,k_3) &= \Gamma_g^{(2)}(\boldsymbol{k}_1,\boldsymbol{k}_2)\,\Gamma_g^{(1)}(k_1)\,\Gamma_g^{(1)}(k_2)\,P_L(k_1)P_L(k_2) + \text{cyc.} \\ &+ \left[\int_{\boldsymbol{q}} \Gamma_g^{(2)}(\boldsymbol{k}_1-\boldsymbol{q},\boldsymbol{q})\,\Gamma_g^{(2)}(\boldsymbol{k}_2+\boldsymbol{q},-\boldsymbol{q})\,\Gamma_g^{(2)}(\boldsymbol{k}_1-\boldsymbol{q},\boldsymbol{k}_2+\boldsymbol{q})\,P_L(|\boldsymbol{k}_1-\boldsymbol{q}|)\,P_L(|\boldsymbol{k}_2+\boldsymbol{q}|)\,P_L(q) \right. \\ &\left. + \frac{1}{2}\int_{\boldsymbol{q}} \Gamma_g^{(3)}(\boldsymbol{k}_3,\boldsymbol{k}_2-\boldsymbol{q},\boldsymbol{q})\,\Gamma_g^{(2)}(\boldsymbol{k}_2-\boldsymbol{q},\boldsymbol{q})\,\Gamma_g^{(1)}(k_3)\,P_L(|\boldsymbol{k}_2-\boldsymbol{q}|)\,P_L(q)\,P_L(k_3) + \text{cyc.} \right]. \end{aligned} \tag{7.96}$$

Therefore, for the one-loop galaxy bispectrum we require both, $\Gamma_g^{(1)}$ and $\Gamma_g^{(2)}$, up to one-loop order in the first term of Eq. (7.96), but tree-level expressions for them and $\Gamma_g^{(3)}$ are enough in the loop integrals, i.e. in the second and third line.

From comparing Eqs. (7.95) and (7.96) we note that the two-point propagator contributes to the leading order bispectrum, while showing up as a loop correction for the power spectrum. This structure extends to consecutively higher orders, for instance, the three-point propagator which appears as a one-loop expression in the bispectrum, will enter at tree-level for the trispectrum. That suggests that constraints on bias parameters required to fit the small-scale behaviour of a given correlator (and thus cosmological parameters, too) will already highly benefit from the large-scale information of the next-order correlator.

7.4.2 Higher-derivative bias and EFT contributions

We now discuss the power spectrum and bispectrum contributions from higher-derivative galaxy bias. We choose to consider them separately from the remaining local and non-local effects because, as we will see below, higher-derivative galaxy bias is mostly degenerate with corrections that have to be taken into account if the dark matter field can no longer be treated as a pressureless perfect fluid. Initially, or on large scales, this is

a good approximation as dark matter particles tend to move within single coherent flows, which implies a vanishing stress tensor σ_{ij}. At later times, however, multi-streaming induces non-zero stresses, which can already have an impact on quasi-linear scales, comparable to the regime where the SPT one-loop terms become important.

The corresponding corrections have been computed in the framework of the effective field theory (EFT) of large-scale structure (Baumann et al., 2012; Carrasco et al., 2012) and as shown in D'Amico and Scoccimarro (2018) can be derived by systematically expanding the stress tensor in powers of k, i.e. considering perturbations at the level of the Vlasov equation. This leads to additional terms in the matter power spectrum and bispectrum that scale as powers of k^2, and at lowest order we have (Baldauf et al., 2015; D'Amico and Scoccimarro, 2018):

$$P_\sigma(k) = -2\beta_P\, k^2\, P_L(k) \tag{7.97}$$

$$B_{\sigma,123} = -\left\{\left[\beta_{B,a}\left(k_1^2+k_2^2\right)+\beta_{B,b}\,k_3^2\right] F_2(\boldsymbol{k}_1,\boldsymbol{k}_2) + \left[\beta_{B,c}\left(k_1^2+k_2^2\right)+\beta_{B,d}\,k_3^2\right] K(\boldsymbol{k}_1,\boldsymbol{k}_2)\right\} \times P_L(k_1)\,P_L(k_2) + \text{cyc.}\,, \tag{7.98}$$

where β_P and $\beta_{B,i}$ are scale independent integrals over stress tensor components and are thus considered to be free parameters. These terms are expected to regularise the effect of the small-scale regime on the loop integrals, which induces the same momentum scaling, as we have seen at the end of Sec. 7.3.3.

On the other hand, by evaluating the first line of Eq. (7.96) using the corrections derived in Eqs. (7.72) and (7.78), we find that the bispectrum contributions due to higher-derivative bias give rise to very similar terms. The ones that are not clearly degenerate with those already appearing in Eq. (7.98) are

$$\text{i)} \quad k_3^2\, P_L(k_1)\, P_L(k_2) + \text{cyc.}\,, \tag{7.99}$$

$$\text{ii)} \quad (\boldsymbol{k}_1\cdot\boldsymbol{k}_2)\, P_L(k_1)\, P_L(k_2) + \text{cyc.}\,, \tag{7.100}$$

$$\text{iii)} \quad (\boldsymbol{k}_1\cdot\boldsymbol{k}_2)\, K(\boldsymbol{k}_1,\boldsymbol{k}_2)\, P_L(k_1)\, P_L(k_2) + \text{cyc.}\,, \tag{7.101}$$

and using that $\boldsymbol{k}_1\cdot\boldsymbol{k}_2 = 1/2\left(k_3^2-k_1^2-k_2^2\right)$ we see that the last term can be written as a combination where $\beta_{B,d} = 1/2 = -\beta_{B,c}$. Furthermore, one can show that

$$k_3^2 = -\left[k_1^2+k_2^2-k_3^2\right] F_2(\boldsymbol{k}_1,\boldsymbol{k}_2) - \left[\frac{5}{7}\left(k_1^2+k_2^2\right)+\frac{2}{7}\,k_3^2\right] K(\boldsymbol{k}_1,\boldsymbol{k}_2)\,, \tag{7.102}$$

such that a combination of all four terms in Eq. (7.98) with $\beta_{B,a}=1$, $\beta_{B,b}=-1$, $\beta_{B,c}=5/7$ and $\beta_{B,d}=2/7$ can also accommodate for Eq. (7.99). Only Eq. (7.100) cannot be expressed through the previous terms and must consequently enter the galaxy bispectrum as an independent contribution. In total, this demonstrates that the EFT contributions to the bispectrum are completely degenerate with those from higher-

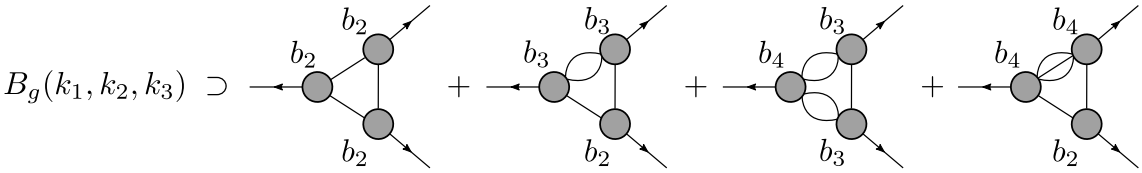


Fig. 7.4 Subset of diagrams that contribute to a non-zero large-scale limit ($k_1, k_2, k_3 \to 0$) of the galaxy bispectrum. The first diagram appears at the one-loop level, the second at two-loop and the last two at three-loop. The bias constants b_i indicate the value of the multi-point propagators (shaded circles).

derivative bias. For that reason we consider them collectively, using the following basis:

$$B_{\sigma+\nabla,123} = -\Big\{\Big[\beta_{B,a}\left(k_1^2+k_2^2\right)+\beta_{B,b}\,k_3^2\Big]\,F_2(\boldsymbol{k}_1,\boldsymbol{k}_2)+\Big[\beta_{B,c}\left(k_1^2+k_2^2\right)+\beta_{B,d}\,k_3^2\Big]\,K(\boldsymbol{k}_1,\boldsymbol{k}_2) \\ +\beta_{B,e}\,\boldsymbol{k}_1\cdot\boldsymbol{k}_2\Big\}\,P_L(k_1)\,P_L(k_2)+\text{cyc.}\,, \tag{7.103}$$

which reduces the number of free parameters to five. Finally, if we restrict ourselves to sufficiently large scales, the EFT and higher-derivative corrections to the power spectrum are degenerate with the Eulerian term proportional to γ_{21} (all three scale as k^2), such that we do not need to include an extra parameter beyond γ_{21}.

7.4.3 Shot noise corrections

The introduction of the multi-point propagators removed unphysical sensitivities to the highly non-linear regime where our perturbative approach is not applicable. However, there are residual effects that become apparent when taking the large-scale limit of the one-loop bispectrum. For instance, let us consider the term that is proportional to b_2^3, deriving from the product of three two-point propagators in the second line of Eq. (7.96):

$$\sim b_2^3\int_{\boldsymbol{q}} P_L(|\boldsymbol{k}_1-\boldsymbol{q}|)\,P_L(|\boldsymbol{k}_2+\boldsymbol{q}|)\,P_L(q) \quad\xrightarrow{k_1,k_2\to 0}\quad b_2^3\int_{\boldsymbol{q}} P_L(q)^3\,. \tag{7.104}$$

We see that in the limit that all three triangle sides approach zero this term does not vanish, but goes to a constant. Such loop corrections therefore affect the bispectrum on asymptotically large scales. Moreover, consecutively higher orders of PT, as shown schematically by the subset of diagrams in Fig. 7.4, contribute to the value of the constant in comparable measures, making it very sensitive to the particular order we are working in. A similar situation arises for the bispectrum loop corrections coming from the last line in Eq. (7.96). Taking one triangle side to zero (from the closure condition the remaining two wave vectors must then be equal and opposite, i.e. a squeezed triangle configuration), one can show that they approach a constant modulated by the power

Table 7.2 Cosmological parameters of the MINERVA simulations. The values of the parameters correspond to the best-fit ΛCDM model of the joint *WMAP9* and *BOSS* Data Release 9 analysis from Sánchez et al. (2013).

Parameter	Ω_m	Ω_Λ	h	n_s	σ_8
Value	0.285	0.715	0.695	0.9632	0.828

spectrum evaluated for the remaining wave vectors. In general,

$$\sim \left[P(k_1) + P(k_2) + P(k_3)\right] \int_{\boldsymbol{q}} P_L(q)^2 \,, \tag{7.105}$$

and this term also receives corrections from increasingly higher orders of PT. We note that Eqs. (7.104) and (7.105) look exactly like the standard Poisson shot noise terms for the bispectrum, so in the spirit of renormalising sensitivities to the small-scale regime by absorbing them into free parameters, we now introduce two *effective* shot noise parameters, such that

$$B_{\rm shot}(k_1, k_2, k_3) = \epsilon_0 + \eta_0 \left[P(k_1) + P(k_2) + P(k_3)\right] . \tag{7.106}$$

We stress that the values of ϵ_0 and η_0 are typically not given by their Poisson predictions, i.e. $\epsilon_0 = 1/\overline{n}^2$ and $\eta_0 = 1/\overline{n}$ for an average number density of galaxies $\overline{n}$, but must be determined from the data itself. The same procedure was already applied to the one-loop power spectrum in McDonald (2006); McDonald and Roy (2009).

As shown in Desjacques et al. (2018), when including higher-derivative corrections, we should also take into account a running of the shot noise parameters above. In case of the bispectrum we then have to introduce three more free parameters as follows:

$$B_{\rm shot,\nabla}(k_1, k_2, k_3) = \epsilon_2\left(k_1^2 + k_2^2 + k_3^2\right) + \left[\eta_{2,1}\, k_1^2 + \eta_{2,2}\left(k_2^2 + k_3^2\right)\right] P(k_1) + \text{cyc.} \tag{7.107}$$

7.5 Application to mock galaxy catalogues

7.5.1 Simulations and measurements

In order to test the performance of the full one-loop bispectrum model, we make use of the MINERVA simulations — a set of 100 N-body simulations that were originally produced in Grieb et al. (2016). Each simulation comprises a total of 1000^3 dark matter particles contained in a cubic box with side length $1500\,h^{-1}$Mpc that was set up with an initial power spectrum at redshift $z_{\rm ini} = 63$ using the cosmological parameters given in Table 7.2. The initial particle distributions were evolved with the N-body code GADGET (Springel, 2005) until redshift $z = 0.57$, which is equivalent with the mean

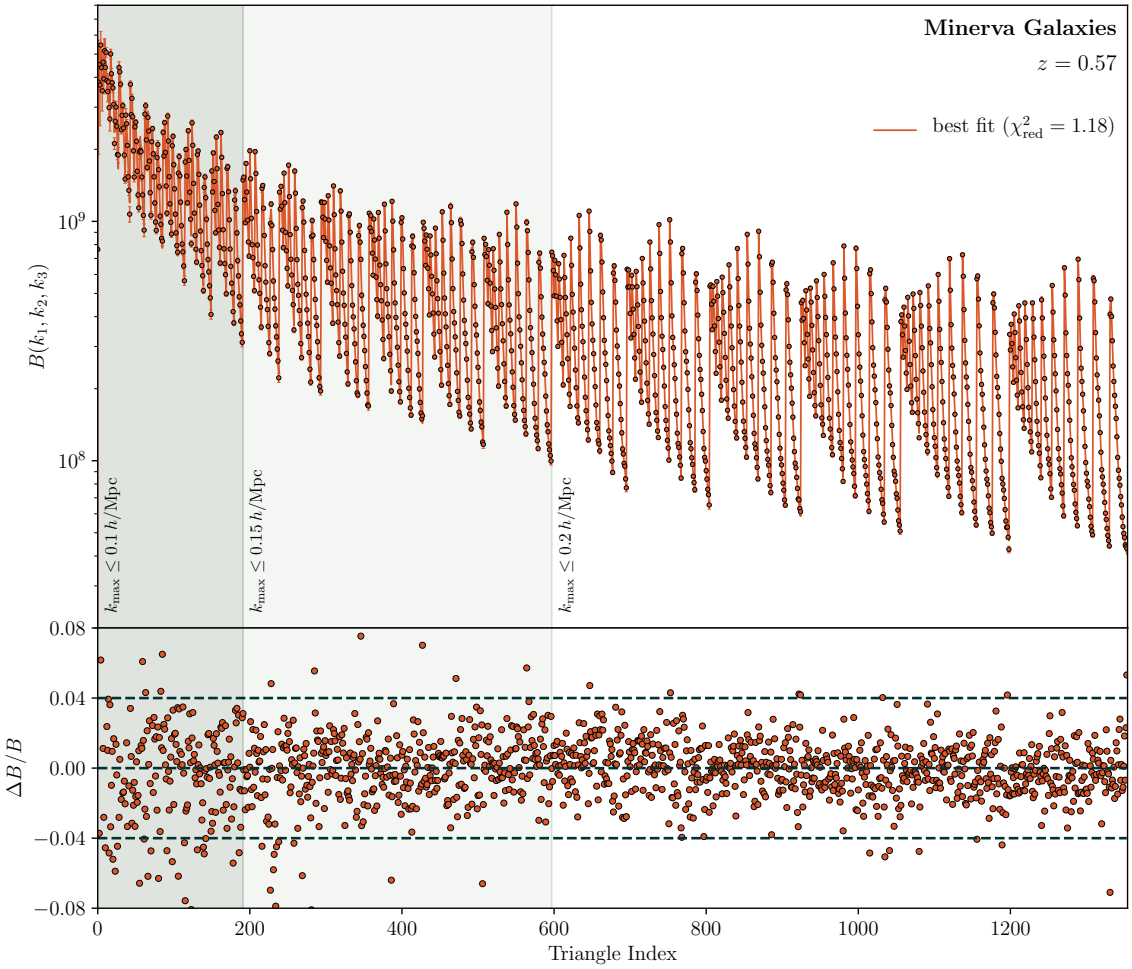


Fig. 7.5 *Top panel:* The measured bispectrum (data points) as a function of triangle index, averaged over the 100 MINERVA simulations. Shaded areas indicate triangle configurations whose three side lengths are smaller than a given maximal value. The solid line represents the one-loop galaxy bispectrum (without higher-derivative corrections) evaluated for the best-fit parameters obtained from a joint power spectrum and bispectrum analysis. *Bottom panel:* Relative difference between the measured and best-fit theory bispectrum.

redshift of the *BOSS* CMASS sample. Subsequently, a friends-of-friends halo finder was applied to generate halo catalogues for each realisation. In a next step, the halos were populated with galaxies following a HOD model whose parameters were tuned such that the resulting galaxy sample has the same two-point clustering properties as the *BOSS* CMASS galaxies (for details, see Grieb et al., 2016). The final galaxy catalogue has a mean number density of $\overline{n} \approx 4 \times 10^{-4}\,h^3\,{\rm Mpc}^{-3}$ and the linear bias, determined by cross-correlating the dark matter and galaxies densities, is given by $b_1 \approx 2.02$.

We have measured both the power spectrum and bispectrum from each galaxy catalogue using an FFT-based estimator (see Chap. 6), and considering a maximum wave mode of $k_{\rm max} = 0.2\,h\,{\rm Mpc}^{-1}$. For the bispectrum we have chosen the spherical bin width $\Delta k = 2k_f \approx 0.0084\,h\,{\rm Mpc}^{-1}$ and looped through all allowed triangle configu-

rations, ordered according to $k_1 \geq k_2 \geq k_3$, which gives a total of 1354 distinct triangle configurations up to $k_1 \leq k_{\rm max}$. We assign each of them a corresponding index, which is used to plot the measured bispectrum, averaged over the 100 simulations, in Fig. 7.5. For guidance we have split the plot into three areas, which indicate triangles whose three sides are smaller than 0.1, 0.15, and 0.2 $h\,\mathrm{Mpc}^{-1}$, respectively.

7.5.2 Constraints on bias parameters

Using the power spectrum and bispectrum data obtained from the mock galaxy catalogues described in the last section, we now perform various model fits. As our principal motivation is to check for consistency between the bias modelling of the power spectrum and bispectrum, we leave the cosmological parameters fixed to the values in Table 7.2. This significantly speeds up the fitting procedure as the bias parameters appear as coefficients in front of the loop integrals, which only need to be computed once if the underlying cosmology does not change. We explore their parameter space by means of the Markov Chain Monte Carlo (MCMC) technique, as implemented by the `PolyChord` package (Handley et al., 2015).

We begin by fitting the vanilla one-loop models without corrections due to higher-derivative terms. Furthermore, we do not explicitly vary the fourth-order bias parameters, as they only enter the bispectrum through a single term (the loop corrections to $\Gamma_g^{(2)}$ in the first line of Eq. 7.96), implying that their impact on the bispectrum will be subdominant. However, we do include the fourth-order bias terms at the level of the local Lagrangian approximation, i.e. their amplitudes are fixed by lower-order bias parameters, because they automatically appear when the multi-point propagators are time evolved from Eq. (7.90). With this approximation we are left with the following 10 parameters in total

$$\text{Bias parameters:} \quad \left\{b_1, b_2, \gamma_2, b_3, \gamma_2^\times, \gamma_3, \gamma_{21}\right\}, \tag{7.108}$$

$$\text{Noise parameters:} \quad \left\{N_0, \epsilon_0, \eta_0\right\}, \tag{7.109}$$

where N_0 is a free shot noise parameter for the power spectrum, which we need to include for the same reason as ϵ_0 and η_0 for the bispectrum (see Sec. 7.4.3).

The model bispectrum evaluated for the best-fit parameters obtained from a joint fit of the power spectrum and bispectrum is plotted in Fig. 7.5. The bottom panel of the figure shows the relative difference between this model prediction and the measured bispectrum, displaying an accuracy better than $\sim 2\,\%$ for the majority of triangle configurations up to $k_{\rm max} = 0.2\,h\,\mathrm{Mpc}^{-1}$. This, and the fact that the reduced chi-square of the joint best-fit result is given by $\chi^2_{\rm red} = 1.18$, indicates consistency between the power spectrum and bispectrum models. This becomes even more apparent if we consider

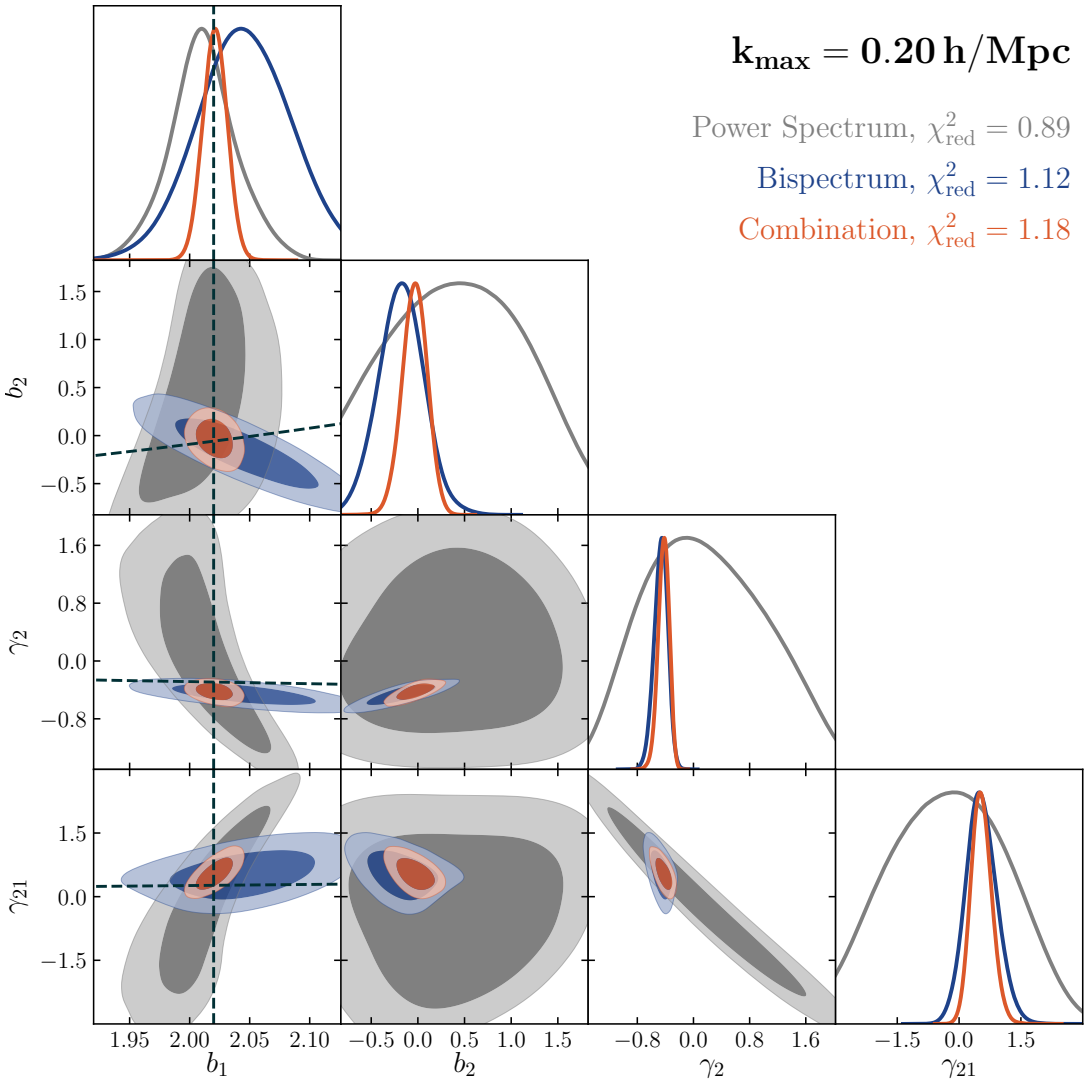


Fig. 7.6 Constraints on a subset of bias parameters. The contours contain 68 % (95 %) of the likelihood when fitting the one-loop power spectrum (grey), one-loop bispectrum (blue), and their combination (orange) up to the maximal scale $k_{max} = 0.2\,h\,\mathrm{Mpc}^{-1}$, excluding higher-derivative terms. The vertical dashed line represents the exact value for the linear bias parameter, whereas the other dashed lines correspond to the PBS and local Lagrangian bias predictions for b_2, γ_2 and γ_{21}, respectively.

the constraints on the bias parameters, a subset of which is shown in Fig. 7.6. The plot shows the 68 % and 95 % likelihood contours for the power spectrum and bispectrum individually (grey and blue contours), as well as their combination (orange). The top panel in each column display the one-dimensional marginalised distribution for the corresponding parameter. We see that the power spectrum and bispectrum are in very good agreement for each combination of parameters, and moreover, the plot demonstrates the power of the bispectrum in reducing uncertainties on bias parameters. This is particularly true for the second-order parameters b_2 and γ_2, which is a direct consequence of them contributing at leading order to the bispectrum, but only at the one-loop level to the power spectrum (see also discussion at the end of Sec. 7.4.1). The

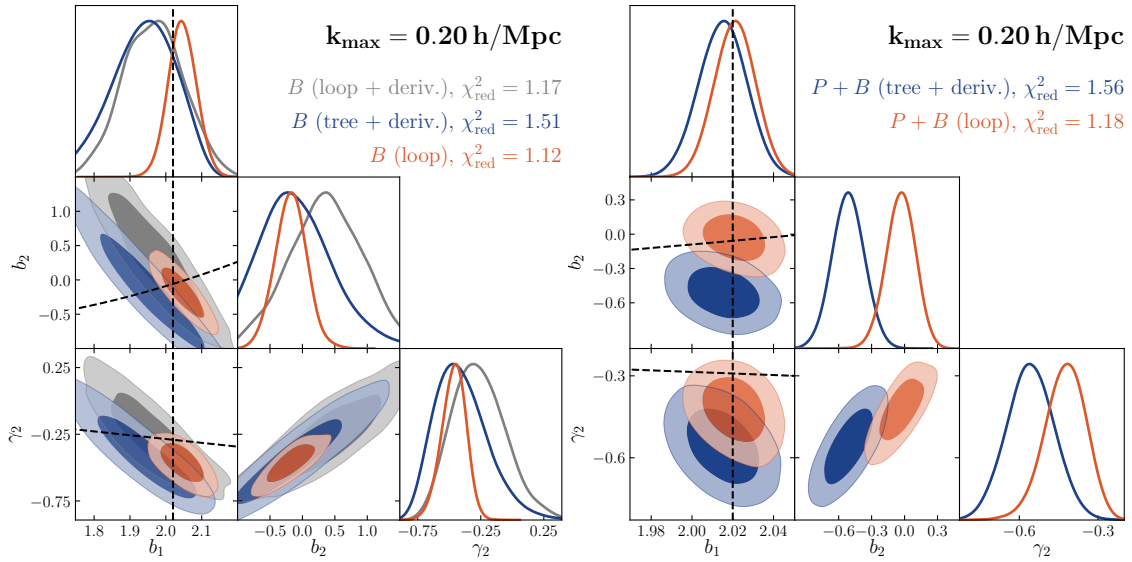


Fig. 7.7 *Left:* Parameter constraints (68 % and 95 % likelihood contours) for the tree-level bias parameters obtained from fitting the bispectrum only. We compare the results using various models: the one-loop bispectrum including higher-derivative terms (grey), tree-level with higher-derivative terms (blue), and the one-loop bispectrum on its own (orange). *Right:* Constraints from the latter two bispectrum models combined with the one-loop power spectrum. Dashed lines have the same meaning as in Fig. 7.6.

third-order parameter γ_{21} is also detected to be non-zero, whereas the remaining three as well as the noise parameters are only weakly constrained. Finally, the vertical dashed line in the first column of Fig. 7.6 represents the exact value of the linear bias parameter, which the joint fit recovers at high accuracy. The remaining dashed lines are predictions from the peak-background split approach (Lazeyras et al., 2016),

$$b_2(b_1) = 0.412 - 2.143\,b_1 + 0.929\,b_1^2 + 0.008\,b_1^3\,, \tag{7.110}$$

and the local Lagrangian values reported in Sec. 7.2.5. Even though the former was strictly speaking derived for dark matter halos, we note that it is in excellent agreement with our results. In general the bias of galaxies will depend on their properties, such as luminosity, colour etc., and therefore differ from the bias of their host dark matter halo. The MINERVA sample, on the other hand, consists of luminous red galaxies that are typically found in the centres of halos, which might explain why they are similarly biased. For the tidal field bias γ_2 we find a value that is slightly below the local Lagrangian approximation, which is consistent with the recent results of Lazeyras and Schmidt (2017).

We now turn to models that include the higher-derivative terms described in Secs. 7.2.4 and 7.4.2. In this case we obtain a set of 8 *additional* parameters (see Eqs. 7.103

and 7.107):

$$\text{Bias parameters:} \quad \{\beta_{B,a}, \beta_{B,b}, \beta_{B,c}, \beta_{B,d}, \beta_{B,e}\}, \tag{7.111}$$

$$\text{Noise parameters:} \quad \{\epsilon_2, \eta_{2,1}, \eta_{2,2}\}, \tag{7.112}$$

following our rationale of not introducing an extra higher-derivative parameter for the power spectrum. The most extensive model — a combination of all one-loop and higher-derivative terms — therefore has a total of 17 free parameters, whereas a simpler version that excludes the loop corrections, requires 8. The constraints on the tree-level bias parameters derived from these two models are shown in the left panel of Fig. 7.7 (represented by the grey and blue contours, respectively), and compared to the previous results (orange contours). We find that the inclusion of higher-derivative terms inflates the uncertainties on the remaining parameters, irrespective of whether the one-loop corrections are taken into account, and even though the tree-level model has less parameters than the ordinary one-loop bispectrum. At the same time the higher-derivative bias parameters themselves remain mostly undetermined. Furthermore, we note a slight shift towards smaller values of b_2 and γ_2 for the tree-level model. This shift becomes more significant when the bispectrum models are combined with the one-loop power spectrum, as demonstrated by the right panel in Fig. 7.7, showing that the constraint on b_2 becomes inconsistent with the peak-background split prediction. These results seem to suggest that: 1) the higher-derivative terms are not required for fitting the bispectrum data up to $k_{\text{max}} = 0.2\,h\,\text{Mpc}^{-1}$ and at redshift $z = 0.57$; and 2) trading the loop corrections for the (computationally easier) higher-derivative terms may introduce systematic biases in the parameter constraints.

7.6 Conclusions

In this chapter we derived a complete basis for galaxy bias based on the equivalence principle and Galilean invariance (see Sec. 7.2.3). We evaluated this basis up to fourth order in perturbation theory, which is required by a one-loop calculation of the bispectrum. Moreover, in Sec. 7.2.2 we presented a more physical way of expanding the galaxy density that converts all bias parameters into measurable quantities, which allows us to circumvent the usual renormalisation problem. From a technical point of view, these measurable quantities are the so-called multi-point propagators, the cross-correlations between the galaxy and matter over-densities. We computed the first three multi-point propagators in Sec. 7.3 based on the previously introduced basis for galaxy bias, and derived a recursion relation that governs their time evolution. Subsequently, we used the multi-point propagators to construct one-loop expressions for the galaxy power spectrum and bispectrum (see Sec. 7.4), and showed that higher-derivative contribu-

tions to the bias basis (induced by the non-locality of galaxy formation) are completely degenerate with corrections expected from the EFT of large-scale structure. In Sec. 7.5 we applied the full one-loop bispectrum model for the first time to data from a set of 100 galaxy mock catalogues, and extracted constraints on the bias parameters by fitting the model using an MCMC technique. We found very good consistency between the one-loop power spectrum and bispectrum up to a scale of $k_{\rm max} = 0.2\,h\,{\rm Mpc}^{-1}$, and demonstrated that the agreement between the best-fit and measured bispectrum is at the level of 2 % for the majority of triangle configurations. Inclusion of higher-derivative terms did not bring about any further improvements — conversely, we found that using the higher-derivative contributions in exchange for the one-loop terms introduces a potential bias in the parameter constraints.

The results presented in Sec. 7.5 are still preliminary and can be extended in several ways. For instance, we can increase the maximum wave number used in the model fits to determine where the overall model breaks down, and whether the inclusion of higher-derivative terms allow us to extend this regime. In general they scale with the Lagrangian radius R_* of the halos and can therefore be expected to be more dominant for galaxies residing in more massive halos. To test their importance it would thus be interesting to extend the present analysis to different samples of galaxies.

Appendices

Appendix 7.A Further notes on Galilean basis for galaxy bias

7.A.1 Basis operators in Fourier space

Here we briefly summarise Fourier space expressions for our basis operators given in Table 7.1, which are being used in the computation of the multi-point propagators. In general, we write any n-th order operator $\mathcal{O}^{(n)}$ as an integral over n linear matter perturbations:

$$\mathcal{O}_B^{(n)}(\boldsymbol{k}) = (2\pi)^3 \int_{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n} \delta_D(\boldsymbol{k}-\boldsymbol{k}_{1\cdots n})\, \mathcal{K}_B^{(n)}(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n) \prod_{i=1}^{n} \delta_L(\boldsymbol{k}_i)\,, \tag{7.113}$$

where B stands for any of the basis operators at that order. For $n=2$ we have from Eq. (7.24):

$$\mathcal{K}_{\delta^2}^{(2)}(\boldsymbol{k}_1,\boldsymbol{k}_2) = \frac{1}{2}\,, \tag{7.114}$$

$$\mathcal{K}_{\mathcal{G}_2}^{(2)}(\boldsymbol{k}_1,\boldsymbol{k}_2) = K(\boldsymbol{k}_1,\boldsymbol{k}_2)\,, \tag{7.115}$$

and the kernel $K(\boldsymbol{k}_1,\boldsymbol{k}_2)$ was already defined in Eq. (7.26). At third order the only nontrivial operator is $\mathcal{G}_2(\varphi_2,\varphi_1)$, which becomes upon Fourier transformation:

$$\begin{aligned}\mathcal{G}_2(\varphi_2,\varphi_1\,|\,\boldsymbol{k}) &= (2\pi)^3 \int_{\boldsymbol{k}_1,\boldsymbol{k}_2} \delta_D(\boldsymbol{k}-\boldsymbol{k}_{12})\, K(\boldsymbol{k}_1,\boldsymbol{k}_2)\, \delta_L(\boldsymbol{k}_1)\, \mathcal{G}_2(\Phi_L\,|\,\boldsymbol{k}_2) \\ &= (2\pi)^3 \int_{\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3} \delta_D(\boldsymbol{k}-\boldsymbol{k}_{123})\, K(\boldsymbol{k}_1,\boldsymbol{k}_{23})\, K(\boldsymbol{k}_2,\boldsymbol{k}_3)\, \delta_L(\boldsymbol{k}_1)\, \delta_L(\boldsymbol{k}_2)\, \delta_L(\boldsymbol{k}_3)\,,\end{aligned} \tag{7.116}$$

where we have made use of Eq. (7.24) and made the redefinition $\boldsymbol{k}_2 \to \boldsymbol{k}_{23}$ in the second step. After symmetrisation we then obtain:

$$\mathcal{K}^{(3)}_{\delta^3}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) = \frac{1}{6}\,, \tag{7.117}$$

$$\mathcal{K}^{(3)}_{\delta\mathcal{G}_2}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) = \frac{1}{3}\Big[K(\boldsymbol{k}_1,\boldsymbol{k}_2) + \text{cyc.}\Big]\,, \tag{7.118}$$

$$\mathcal{K}^{(3)}_{\mathcal{G}_3}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) = L(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)\,, \tag{7.119}$$

$$\mathcal{K}^{(3)}_{\mathcal{G}_2(\varphi_2,\varphi_1)}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) = \frac{1}{3}\Big[K(\boldsymbol{k}_1,\boldsymbol{k}_{23})\,K(\boldsymbol{k}_2,\boldsymbol{k}_3) + \text{cyc.}\Big]\,. \tag{7.120}$$

Next, let us consider the most complicated combination that appears at fourth order, $\mathcal{G}_2(\varphi_3,\varphi_1)$, all other operators will follow in a very similar manner. Starting from the definition in Eq. (7.41) and using the relations (7.38) to (7.40) for the LPT potentials, we have

$$\begin{aligned}\mathcal{G}_2(\varphi_3,\varphi_1\,|\,\boldsymbol{k}) = (2\pi)^3 \int\limits_{\boldsymbol{k}_1,\boldsymbol{k}_2} \delta_D(\boldsymbol{k}-\boldsymbol{k}_{12}) \Bigg[\frac{1}{18}K(\boldsymbol{k}_1,\boldsymbol{k}_2)\left(\mathcal{G}_3(\varphi_1\,|\,\boldsymbol{k}_2) + \frac{15}{7}\mathcal{G}_2(\varphi_2,\varphi_1\,|\,\boldsymbol{k}_2)\right)\\ + \frac{1}{14}\frac{(\boldsymbol{k}_1\cdot\boldsymbol{k}_2)\;k_{1,j}\,k_{2,l}}{k_1^2\,k_2^2}\left[\nabla_{lm}\varphi_2\,\nabla_{jm}\varphi_1 - \nabla_{jm}\varphi_2\,\nabla_{lm}\varphi_1\right](\boldsymbol{k}_2)\Bigg]\delta_L(\boldsymbol{k}_1)\,.\end{aligned} \tag{7.121}$$

Plugging in the Fourier expressions for the remaining potentials and Galileons (using Eq. 7.116), and replacing $\boldsymbol{k}_2 \to \boldsymbol{k}_{234}$ we get

$$\begin{aligned}&\mathcal{G}_2(\varphi_3,\varphi_1\,|\,\boldsymbol{k})\\ &= (2\pi)^3 \int\limits_{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_4} \delta_D(\boldsymbol{k}-\boldsymbol{k}_{1234}) \Bigg[\frac{1}{18}K(\boldsymbol{k}_1,\boldsymbol{k}_{234})\left(\frac{15}{7}K(\boldsymbol{k}_{23},\boldsymbol{k}_4)\,K(\boldsymbol{k}_2,\boldsymbol{k}_3) - L(\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4)\right)\\ &+ \frac{1}{14}\Big(M(\boldsymbol{k}_1,\boldsymbol{k}_{234},\boldsymbol{k}_{23},\boldsymbol{k}_4) - M(\boldsymbol{k}_1,\boldsymbol{k}_{23},\boldsymbol{k}_4,\boldsymbol{k}_{234})\Big)K(\boldsymbol{k}_2,\boldsymbol{k}_3)\Bigg]\prod_{i=1}^{4}\delta_L(\boldsymbol{k}_i)\,,\end{aligned} \tag{7.122}$$

where we have introduced the new kernel

$$M(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4) \equiv \frac{(\boldsymbol{k}_1\cdot\boldsymbol{k}_2)\,(\boldsymbol{k}_2\cdot\boldsymbol{k}_3)\,(\boldsymbol{k}_3\cdot\boldsymbol{k}_4)\,(\boldsymbol{k}_4\cdot\boldsymbol{k}_1)}{(k_1\,k_2\,k_3\,k_4)^2}\,, \tag{7.123}$$

which is symmetric under cyclic permutations of its four momenta. The fully symmetric kernels for the basis operators at fourth order are thus given by

$$\mathcal{K}^{(3)}_{\delta^4}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4) \equiv \frac{1}{24}\,, \tag{7.124}$$

$$\mathcal{K}^{(3)}_{\delta^2\mathcal{G}_2}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4) \equiv \frac{1}{6}\Big[K(\boldsymbol{k}_1,\boldsymbol{k}_2)+\text{sym.}\Big]\,, \tag{7.125}$$

$$\mathcal{K}^{(3)}_{\delta\mathcal{G}_3}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4) \equiv \frac{1}{4}\Big[L(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)+\text{cyc.}\Big]\,, \tag{7.126}$$

$$\mathcal{K}^{(3)}_{\mathcal{G}_2^2}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4) \equiv \frac{1}{3}\Big[K(\boldsymbol{k}_1,\boldsymbol{k}_2)\,K(\boldsymbol{k}_3,\boldsymbol{k}_4)+\text{sym.}\Big]\,, \tag{7.127}$$

$$\mathcal{K}^{(3)}_{\delta\mathcal{G}_2(\varphi_2,\varphi_1)}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4) \equiv \frac{1}{12}\Big[K(\boldsymbol{k}_1,\boldsymbol{k}_{23})\,K(\boldsymbol{k}_2,\boldsymbol{k}_3)+\text{sym.}\Big]\,, \tag{7.128}$$

$$\mathcal{K}^{(3)}_{\mathcal{G}_3(\varphi_2,\varphi_1,\varphi_1)}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4) \equiv \frac{1}{6}\Big[L(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_{34})\,K(\boldsymbol{k}_3,\boldsymbol{k}_4)+\text{sym.}\Big]\,, \tag{7.129}$$

$$\mathcal{K}^{(3)}_{\mathcal{G}_2(\varphi_2,\varphi_2)}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4) \equiv \frac{1}{3}\Big[K(\boldsymbol{k}_{12},\boldsymbol{k}_{34})\,K(\boldsymbol{k}_1,\boldsymbol{k}_2)\,K(\boldsymbol{k}_3,\boldsymbol{k}_4)+\text{sym.}\Big]\,, \tag{7.130}$$

$$\begin{aligned}\mathcal{K}^{(3)}_{\mathcal{G}_2(\varphi_3,\varphi_1)}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4) \equiv \frac{1}{72}\Bigg[&K(\boldsymbol{k}_1,\boldsymbol{k}_{234})\left(\frac{5}{7}\,K(\boldsymbol{k}_{23},\boldsymbol{k}_4)\,K(\boldsymbol{k}_2,\boldsymbol{k}_3)-L(\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4)\right)\\ &+\frac{3}{7}\Big(M(\boldsymbol{k}_1,\boldsymbol{k}_{234},\boldsymbol{k}_{23},\boldsymbol{k}_4)-M(\boldsymbol{k}_1,\boldsymbol{k}_{23},\boldsymbol{k}_4,\boldsymbol{k}_{234})\Big)K(\boldsymbol{k}_2,\boldsymbol{k}_3)+\text{sym.}\Bigg]\,.\end{aligned} \tag{7.131}$$

7.A.2 Relation to bias basis of Mirbabayi et al.

The basis for galaxy bias given in Desjacques et al. (2018); Mirbabayi et al. (2015) is expressed in terms of $\Pi^{[1]}_{ij}(\boldsymbol{x},\tau) \equiv \nabla_{ij}\Phi(\boldsymbol{x},\tau)$ and its convective time derivatives. Using these ingredients they recursively define the tensors

$$\Pi^{[n]}_{ij} \equiv \frac{1}{(n-1)!}\left[\frac{1}{\mathscr{H}f}\frac{\mathrm{d}}{\mathrm{d}\tau}\Pi^{[n-1]}_{ij}-(n-1)\,\Pi^{[n-1]}_{ij}\right]\,, \tag{7.132}$$

whose leading order contributions come from n-th order in PT ($\mathrm{d}/\mathrm{d}\tau$ denotes the convective time derivative, i.e. $\mathrm{d}/\mathrm{d}\tau \equiv \partial/\partial\tau + \boldsymbol{v}\cdot\nabla$). By combining various tensors and computing their scalar contractions, (Desjacques et al., 2018; Mirbabayi et al., 2015) show how to construct a set of linearly independent operators at each order of PT. The resulting basis is completely equivalent with the one introduced in Sec. 7.2.3, and in the following we give explicit expressions that relate the two.

At linear order it is only possible to take the trace of $\Pi^{[1]}_{ij}$, which is simply

$$\mathrm{Tr}\left[\Pi^{[1]}\right] = \delta\,. \tag{7.133}$$

Apart from the square of Eq. (7.133), there is one more operator at second order, which is related to our operators as

$$\mathrm{Tr}\left[\left(\Pi^{[1]}\right)^2\right] = \delta^2 + \mathscr{G}_2\,. \tag{7.134}$$

At third order there are two additional operators, which are not products of lower order ones, and for these we find:

$$\mathrm{Tr}\left[\left(\Pi^{[1]}\right)^3\right] = \delta^3 + \frac{3}{2}\delta\mathscr{G}_2 + \frac{1}{2}\mathscr{G}_3\,, \tag{7.135}$$

$$\mathrm{Tr}\left[\Pi^{[1]}\,\Pi^{[2]}\right] = \delta^3 + \frac{11}{14}\delta\mathscr{G}_2 + \frac{1}{2}\mathscr{G}_3 - \frac{5}{7}\mathscr{G}_2(\varphi_2,\varphi_1)\,. \tag{7.136}$$

Finally, the fourth order basis requires us to map a total of four operators, which can be shown to satisfy the relations,

$$\mathrm{Tr}\left[\left(\Pi^{[1]}\right)^4\right] = \delta^4 + \frac{1}{2}\mathscr{G}_2^2 + 2\delta^2\mathscr{G}_2 + \frac{2}{3}\delta\mathscr{G}_3\,, \tag{7.137}$$

$$\begin{aligned}\mathrm{Tr}\left[\Pi^{[2]}\,\Pi^{[2]}\right] = \delta^4 &+ \frac{29}{98}\mathscr{G}_2^2 + \frac{4}{7}\delta^2\mathscr{G}_2 + \frac{2}{3}\delta\mathscr{G}_3 - \frac{10}{7}\delta\mathscr{G}_2(\varphi_2,\varphi_1)\\ &+ \frac{25}{49}\mathscr{G}_2(\varphi_2,\varphi_2) - \frac{5}{7}\mathscr{G}_3(\varphi_2,\varphi_1,\varphi_1)\,,\end{aligned} \tag{7.138}$$

$$\mathrm{Tr}\left[\Pi^{[1]}\,\Pi^{[1]}\,\Pi^{[2]}\right] = \delta^4 + \frac{1}{7}\mathscr{G}_2^2 + \frac{9}{7}\delta^2\mathscr{G}_2 + \frac{2}{3}\delta\mathscr{G}_3 - \frac{5}{7}\delta\mathscr{G}_2(\varphi_2,\varphi_1) - \frac{5}{14}\mathscr{G}_3(\varphi_2,\varphi_1,\varphi_1)\,, \tag{7.139}$$

$$\begin{aligned}\mathrm{Tr}\left[\Pi^{[1]}\,\Pi^{[3]}\right] = 2\delta^4 &+ \frac{1}{14}\mathscr{G}_2^2 + \frac{15}{7}\delta^2\mathscr{G}_2 + \frac{5}{9}\delta\mathscr{G}_3 - \frac{4}{21}\delta\mathscr{G}_2(\varphi_2,\varphi_1)\\ &+ \frac{13}{14}\mathscr{G}_3(\varphi_2,\varphi_1,\varphi_1) + 14\mathscr{G}_2(\varphi_3,\varphi_1)\,.\end{aligned} \tag{7.140}$$

As is evident from the relations above, the basis by Desjacques et al. (2018); Mirbabayi et al. (2015) mixes our Lagrangian and Eulerian operators.

Appendix 7.B Time evolution of multi-point propagators

In this appendix we give a detailed derivation for the time evolution of the multi-point propagators. We proceed in two steps: first, we evaluate expectation values of products of two or three Wiener-Hermite functionals, and second, by using these results we directly show how to obtain the recursion relations reported in Eq. (7.91).

7.B.1 Orthogonality relations for generalised Wiener-Hermite functionals

Let us consider the PDF of δ_L, shifted by a generic source term $\alpha(\boldsymbol{k})$, which we take to be an arbitrary function of wave number $\boldsymbol{k}$. A Taylor expansion around $\alpha = 0$ yields:

$$\mathscr{P}[\alpha - \delta_L] = \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n} \left.\frac{\partial^n \mathscr{P}[\alpha - \delta_L]}{\partial\alpha_1 \cdots \partial\alpha_n}\right|_{\alpha=0} \alpha_1 \cdots \alpha_n \,, \tag{7.141}$$

where $\alpha_i \equiv \alpha(\boldsymbol{k}_i)$. Swapping the derivatives from α to δ_L and using the definition of the Wiener-Hermite functionals from Eq. (7.57) we get

$$\left.\frac{\partial^n \mathscr{P}[\alpha - \delta_L]}{\partial\alpha_1 \cdots \partial\alpha_n}\right|_{\alpha=0} = (-1)^n \left.\frac{\partial^n \mathscr{P}[\delta_L]}{\partial\delta_{L,1} \cdots \partial\delta_{L,n}}\right|_{\alpha=0} = \frac{\mathscr{P}[\delta_L]\, \mathscr{H}_n(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n)}{P_L(k_1)\cdots P_L(k_n)} \,, \tag{7.142}$$

and thus:

$$\frac{\mathscr{P}[\alpha - \delta_L]}{\mathscr{P}[\delta_L]} = \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n} \frac{\mathscr{H}_n(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n)}{P_L(k_1)\cdots P_L(k_n)} \alpha_1 \cdots \alpha_n \,. \tag{7.143}$$

To derive the orthogonality relation between two Wiener-Hermite functionals of orders m and n, we first compute the following integral

$$\int \mathscr{D}[\delta_L]\, \mathscr{P}[\delta_L]\, \frac{\mathscr{P}[\alpha - \delta_L]}{\mathscr{P}[\delta_L]}\, \frac{\mathscr{P}[\beta - \delta_L]}{\mathscr{P}[\delta_L]} = \exp\left[\int_{\boldsymbol{q}} \frac{\alpha(\boldsymbol{q})\,\beta(-\boldsymbol{q})}{P_L(q)}\right], \tag{7.144}$$

where we have plugged in Eq. (7.56). However, using Eq. (7.143) to replace the PDF's, we must also have:

$$\begin{aligned} \sum_{m,n} \frac{1}{m!\,n!} \int_{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_m} \int_{\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n} \frac{\langle \mathscr{H}_m(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_m)\, \mathscr{H}_n(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n)\rangle}{P_L(k_1)\cdots P_L(k_m)\, P_L(q_1)\cdots P_L(q_n)} \alpha_1 \cdots \alpha_m \beta_1 \cdots \beta_n \\ = \exp\left[\int_{\boldsymbol{q}} \frac{\alpha(\boldsymbol{q})\,\beta(-\boldsymbol{q})}{P_L(q)}\right], \end{aligned} \tag{7.145}$$

and by Taylor expanding the right-hand side of the expression above, we see that we need to match up all $\boldsymbol{k}$ and $\boldsymbol{q}$ modes, which is only possible if $m = n$. From that observation it immediately follows that

$$\langle \mathscr{H}_m(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_m)\, \mathscr{H}_n(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n)\rangle = (2\pi)^{3m} \delta_{mn}\, \delta_D\left(\boldsymbol{k}_{\{1,m\}}, \boldsymbol{q}_{\{1,m\}}\right) \prod_{i=1}^{m} P_L(k_i) \,, \tag{7.146}$$

where we have used the short-hand notation,

$$\delta_D\left(\boldsymbol{k}_{\{1,m\}}, \boldsymbol{q}_{\{1,m\}}\right) \equiv \delta_D(\boldsymbol{k}_1 + \boldsymbol{q}_1) \cdots \delta_D(\boldsymbol{k}_m + \boldsymbol{q}_m) + \text{sym.} \,, \tag{7.147}$$

for writing all possible ways ($m!$ in total) of matching up the two sets of modes.

Let us now compute the expectation value of three Wiener-Hermite functionals, i.e. $\langle \mathscr{H}_m \mathscr{H}_n \mathscr{H}_l \rangle$. Similar to the above procedure, we first evaluate an integral over PDF's, now with the three different sources α, β and γ:

$$\begin{aligned} &\int \mathscr{D}[\delta_L]\, \mathscr{P}[\delta_L]\, \frac{\mathscr{P}[\alpha-\delta_L]}{\mathscr{P}[\delta_L]}\, \frac{\mathscr{P}[\beta-\delta_L]}{\mathscr{P}[\delta_L]}\, \frac{\mathscr{P}[\gamma-\delta_L]}{\mathscr{P}[\delta_L]} \\ &= \exp\left[\int_{\boldsymbol{q}} \frac{\alpha(\boldsymbol{q})\,\beta(-\boldsymbol{q}) + \alpha(\boldsymbol{q})\,\gamma(-\boldsymbol{q}) + \beta(\boldsymbol{q})\,\gamma(-\boldsymbol{q})}{P_L(q)} \right]. \end{aligned} \tag{7.148}$$

Expanding both sides of Eq. (7.148), we get

$$\begin{aligned} (7.148, \text{left}) = \sum_{m,n,l} \frac{1}{m!\,n!\,l!} \int_{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_m} \int_{\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n} \int_{\boldsymbol{p}_1,\ldots,\boldsymbol{p}_l} \left[\prod_{i=1}^{m} \frac{\alpha(\boldsymbol{k}_i)}{P_L(k_i)}\right] \left[\prod_{i=1}^{n} \frac{\beta(\boldsymbol{q}_i)}{P_L(q_i)}\right] \left[\prod_{i=1}^{l} \frac{\gamma(\boldsymbol{p}_i)}{P_L(p_i)}\right] \\ \times \langle \mathscr{H}_m(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_m)\, \mathscr{H}_n(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n)\, \mathscr{H}_l(\boldsymbol{p}_1,\ldots,\boldsymbol{p}_l) \rangle\,, \end{aligned} \tag{7.149}$$

and

$$\begin{aligned} (7.148, \text{right}) = \sum_{a,b,c} \frac{1}{a!\,b!\,c!} \int_{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_a} \int_{\boldsymbol{q}_1,\ldots,\boldsymbol{q}_b} \int_{\boldsymbol{p}_1,\ldots,\boldsymbol{p}_c} \left[\prod_{i=1}^{a} \frac{\alpha(\boldsymbol{k}_i)\,\beta(-\boldsymbol{k}_i)}{P_L(k_i)}\right] \\ \times \left[\prod_{i=1}^{b} \frac{\alpha(\boldsymbol{q}_i)\,\gamma(-\boldsymbol{q}_i)}{P_L(q_i)}\right] \left[\prod_{i=1}^{c} \frac{\beta(\boldsymbol{p}_i)\,\gamma(-\boldsymbol{p}_i)}{P_L(p_i)}\right]. \end{aligned} \tag{7.150}$$

In order for Eqs. (7.149) and (7.150) to be equal, they need to contain the same number of source terms, which requires that $l+m+n = 2(a+b+c)$ and thus, $l+m+n \in 2\mathbb{N}$. Moreover, the indices must satisfy the conditions

$$\left.\begin{aligned} a+b &= m \\ a+c &= n \\ b+c &= l \end{aligned}\right\} \Leftrightarrow \left\{\begin{aligned} a &= \frac{m+n-l}{2} \\ b &= \frac{l+m-n}{2} \\ c &= \frac{n+l-m}{2} \end{aligned}\right., \tag{7.151}$$

from which follows that $m+n \geq l$, and cyclic permutations thereof. According to Eq. (7.151), we can divide all $\boldsymbol{k}$-, $\boldsymbol{q}$-, and $\boldsymbol{p}$-modes into two subsets each, either of size a, b or c. For instance, the $\boldsymbol{k}$-modes will be split into a group containing a modes, and another containing b, such that $a+b=m$. Each mode in the former group can then be assigned a mode from an equally sized group of $\boldsymbol{q}$'s, i.e. $\boldsymbol{q}_1 = -\boldsymbol{k}_1, \ldots, \boldsymbol{q}_a = -\boldsymbol{k}_a$, while the latter are matched in a similar manner with a set of $\boldsymbol{p}$-modes, which guarantees that we obtain Eq. (7.150). Altogether, this means we need to require that the expectation

value of the three Wiener-Hermite functionals is given by

$$\begin{aligned}\langle \mathscr{H}_m(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_m)\,\mathscr{H}_n(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n)\,\mathscr{H}_l(\boldsymbol{p}_1,\ldots,\boldsymbol{p}_l)\rangle \\ = (2\pi)^{\frac{3}{2}(m+n+l)}\left[\prod_{i=1}^{b} P_L(k_{a+i})\right]\left[\prod_{i=1}^{a} P_L(q_i)\right]\left[\prod_{i=1}^{c} P_L(p_{b+i})\right] \\ \times\left[\delta_D\left(\boldsymbol{k}_{\{1,a\}},\boldsymbol{q}_{\{1,a\}}\right)\delta_D\left(\boldsymbol{k}_{\{a+1,m\}},\boldsymbol{p}_{\{1,b\}}\right)\delta_D\left(\boldsymbol{q}_{\{a+1,n\}},\boldsymbol{p}_{\{b+1,l\}}\right)+\text{sym.}\right],\end{aligned} \tag{7.152}$$

and it must vanish if the conditions above are not satisfied. The number of terms that have to be added in order to symmetrise Eq. (7.152) is given by the number of possibilities of selecting subsets of $\boldsymbol{k}$-, $\boldsymbol{q}$-, and $\boldsymbol{p}$-modes that are of size b, a and c, respectively. The total number of terms in the square brackets is thus:

$$\binom{m}{b}\times\binom{n}{a}\times\binom{l}{c}\times a!\,b!\,c! = \frac{m!\,n!\,l!}{a!\,b!\,c!}\,. \tag{7.153}$$

7.B.2 The Γ-recursion relation

We are interested in the time evolution of the n-th order multi-point propagator. Exploiting the orthogonality of Wiener-Hermite functionals, we can single out its contribution to the series expansion in Eq. (7.89) by multiplying both sides with $\mathscr{H}_n$ and taking the ensemble average:

$$\langle \mathscr{H}_n(\delta_L\,|\,\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n)\,\Psi_a(\boldsymbol{k},\eta)\rangle = (2\pi)^3\,\delta_D(\boldsymbol{k}-\boldsymbol{k}_{1\cdots n})\,\Gamma_a^{(n)}(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n;\eta)\prod_{i=1}^{n} P_L(k_i) \tag{7.154}$$

Applying the same procedure to the integral solution for Ψ_a from Eq. (7.87) we get

$$\begin{aligned}\Gamma_a^{(n)}(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n;\eta) = g_{ab}(\eta)\,\widehat{\Gamma}_b^{(n)}(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n) + \sum_{m,l=1}\frac{(2\pi)^3}{m!\,l!}\int_0^\eta \mathrm{d}\eta'\,g_{ab}(\eta-\eta') \\ \times\int_{\boldsymbol{q}_1,\ldots,\boldsymbol{q}_m}\int_{\boldsymbol{p}_1,\ldots,\boldsymbol{p}_l}\gamma_{bcd}(\boldsymbol{k}_{1\cdots n},\boldsymbol{q}_{1\cdots m},\boldsymbol{p}_{1\cdots l})\,\Gamma_c^{(m)}(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_m;\eta')\,\Gamma_d^{(l)}(\boldsymbol{p}_1,\ldots,\boldsymbol{p}_l;\eta') \\ \times\langle \mathscr{H}_n(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n)\mathscr{H}_m^*(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_m)\mathscr{H}_l^*(\boldsymbol{p}_1,\ldots,\boldsymbol{p}_l)\rangle\prod_{i=1}^{n}\frac{1}{P_L(k_i)}\,,\end{aligned} \tag{7.155}$$

where $\widehat{\Gamma}_a$ denotes multi-point propagators at initial time $\eta = 0$. Upon inserting Eq. (7.152) we perform the first $a = (n+m-l)/2$ integrations over $\boldsymbol{q}$-modes as well as over all of

the $\boldsymbol{p}$-modes, such that the second term above reduces to

$$\sum_{\substack{m,\,l=1\\ n+m+l\in 2\mathbb{N}\\ n+m\geq l,\ \text{cyc.}}} \frac{(2\pi)^3}{m!\,l!}\,\kappa \int_0^\eta \mathrm{d}\eta'\, g_{ab}(\eta-\eta') \int\limits_{\boldsymbol{q}_{a+1},\ldots,\boldsymbol{q}_m} \Big[\gamma_{bcd}(\boldsymbol{k}_{1\cdots n}, \boldsymbol{k}_{1\cdots a}+\boldsymbol{q}_{a+1\cdots m}, \boldsymbol{k}_{a+1\cdots n}-\boldsymbol{q}_{a+1\cdots m})$$
$$\times\, \Gamma_c^{(m)}(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_a,\boldsymbol{q}_{a+1},\ldots,\boldsymbol{q}_m;\eta')\,\Gamma_d^{(l)}(\boldsymbol{k}_{a+1},\ldots,\boldsymbol{k}_n,-\boldsymbol{q}_{a+1},\ldots,-\boldsymbol{q}_m;\eta') + \text{sym.}\Big]$$
$$\times \prod_{i=a+1}^{m} P_L(q_i)\,. \tag{7.156}$$

The symmetrisation is carried out by summing over all $\binom{n}{a}$ subsets of $\boldsymbol{k}$-modes, which implies that the combinatorial factor κ is given by:

$$\kappa = \frac{m!\,n!\,l!}{a!\,b!\,c!}\binom{n}{a}^{-1} = \frac{m!\,l!}{(m-a)!}\,. \tag{7.157}$$

Next, let us change the summation index from l to $r=(l+m-n)/2=m-a$ and relabel the remaining integrations over $\boldsymbol{q}$'s into $\boldsymbol{q}_1,\ldots,\boldsymbol{q}_r$. The conditions $n+m\geq l$ (and cyclic permutations) then transform into $r_*\leq r\leq m$, where

$$r_* = \begin{cases} 0\,, & n>m \\ 1\,, & n=m \\ m-n\,, & n<m \end{cases}\,. \tag{7.158}$$

Defining the quantity

$$\Gamma_a^{(n,m,r)}(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n;\eta) \equiv \int\limits_{\boldsymbol{q}_1,\ldots,\boldsymbol{q}_r} \Big[\gamma_{abc}(\boldsymbol{k}_{1\cdots n}, \boldsymbol{k}_{1\cdots m-r}+\boldsymbol{q}_{1\cdots r}, \boldsymbol{k}_{m-r+1\cdots n}-\boldsymbol{q}_{1\cdots r})$$
$$\times\, \Gamma_c^{(2r+n-m)}(\boldsymbol{k}_{m-r+1},\ldots,\boldsymbol{k}_n,-\boldsymbol{q}_1,\ldots,-\boldsymbol{q}_r;\eta)$$
$$\times\, \Gamma_b^{(m)}(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_{m-r},\boldsymbol{q}_1,\ldots,\boldsymbol{q}_r;\eta)\prod_{i=1}^{r} P_L(\boldsymbol{q}_i) + \text{sym.}(\boldsymbol{k}_i)\Big]\,, \tag{7.159}$$

we finally obtain the desired recursion relation for multi-point propagators:

$$\Gamma_a^{(n)}(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n;\eta) = g_{ab}(\eta)\,\widehat{\Gamma}_b^{(n)}(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n)$$
$$+ \sum_{m=1}^{m}\sum_{r=r_*}^{m} \frac{(2\pi)^3}{r!}\int_0^\eta \mathrm{d}\eta'\, g_{ab}(\eta-\eta')\,\Gamma_b^{(n,m,r)}(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n;\eta')\,. \tag{7.160}$$

Contributions

The content presented in this chapter originates from a collaborative project with R. Scoccimarro and my supervisor, R. E. Smith, and will soon be published as a paper. I have carried out the great majority of the computations in this chapter, but profited greatly from the assistance of R. Scoccimarro, who provided ideas (the multi-point propagator formalism and the basis for galaxy bias) and directions. Furthermore, I have implemented the complete one-loop galaxy bispectrum, building on an existing code for the power spectrum (written by R. Scoccimarro and collaborators), and performed all model fits. The measurements of the power spectrum and bispectrum were provided by R. Scoccimarro. I wrote the entire manuscript with the exception of Sec. 7.2.2, and produced all of the figures.

Chapter 8

Conclusions

This thesis has focused on three-point statistics for characterising the large-scale distribution of galaxies in our Universe. As they can distinguish between a range of different shapes, three-point statistics are a far richer probe of galaxy clustering than the traditional two-point correlations. For that reason they provide a considerable amount of additional information (see Fig. 3.5) that can be used to test our current standard model of cosmology (ΛCDM, introduced in Sec. 1.1). Despite being very successful, this model predicts that 95 % of our Universe is composed of unknown substances: dark matter and dark energy. Pinning down their nature is the primary goal of many experiments that are going to be conducted within the next decade (Sec. 3.3.4). Three-point statistics can contribute to this goal as they enable a precise measurement of the growth rate of structures by breaking degeneracies between various model parameters (Secs. 3.2.2 and 3.3.3). Such a measurement can tell us about the possible breakdown of general relativity on very large scales, which is currently one of the discussed explanations for dark energy (Sec. 1.2).

However, a reliable extraction of information from three-point statistics, such as the bispectrum, requires a robust and unbiased prediction of its signal for a given cosmological model, as well as its measurement errors. Both come with additional challenges compared to the two-point case, the former because it involves higher orders of perturbation theory (Secs. 2.2.3 and 3.2), and the latter because of the increased size of the covariance matrix (Sec. 3.3.2). We have addressed both of these challenges in this thesis and have achieved two important milestones that will pave the way for a future application of the bispectrum: (1) we have identified a method to shrink the dimension of the covariance matrix by at least an order of magnitude; and (2) we have developed and tested an improved galaxy bias model for the bispectrum that is consistent with state-of-the-art models for two-point statistics. A more detailed summary of our findings is given in Sec. 8.1 and in Sec. 8.2 we sketch a few steps for the road ahead.

8.1 Summary of results

We began our investigations with a particular three-point statistic that is only sensitive to information in the Fourier phases of the density field: the line correlation function (LCF). By definition, the LCF measures information that are complementary to two-point statistics as these are blind to the phases of Fourier modes (Sec. 3.1.3). On large scales the phases of the matter and galaxy distribution are expected to be the same, which makes the LCF potentially a direct probe of the growth rate of structures, i.e. it is not degenerate with the linear bias parameter as is the case for the power spectrum (see Eq. 3.40). The growth rate can be detected from redshift space distortions (Sec. 3.3.3), which lead to an anisotropy in the observed distribution of galaxies. In order to pick up such a signal, it is necessary to account for the line-of-sight and transverse distances between galaxies separately, which motivated the generalisation of the LCF presented in Chap. 4. An integral part in the definition of the LCF is an isotropic high frequency mode cut-off (in order to suppress noise from uncorrelated phases), and we found that leaving it spherical tends to smear out any anisotropic signal. Therefore, we proposed an aspherical cut-off, which improves the signal-to-noise of the LCF by up to an order of magnitude and was shown to be sensitive to angularly varying number densities and scales of structures. We further demonstrated that our generalisation of the LCF is capable of identifying the AP effect and redshift space distortions, albeit at a slightly smaller significance than the two-point function. Our prediction of a change of sign in the quadrupole of the redshift space LCF (see Fig. 4.10) has by now been confirmed by Franco et al. (2018) using non-linear perturbation theory.

In order to gauge the usefulness of the LCF for cosmology, we need to assess its sensitivity on cosmological parameters. The foundation for that was laid in Chap. 5 using two complementary approaches. One approach was based on perturbation theory, yielding a more realistic prediction of the LCF including a non-linear model for galaxy bias and shot noise due to a finite number of galaxies, but for simplicity ignoring the complications from redshift space distortions. We also computed its covariance properties, showing that it is independent of the adopted cosmological model in the Gaussian limit. The second approach was driven numerically, where we measured the LCF from a set of 200 N-body simulations at various redshifts. That allowed us to obtain precise estimates of the LCF amplitude and covariance matrix, finding very good agreement upon comparing them with our theoretical predictions (cf. Figs. 5.3 to 5.5). These ingredients were subsequently used to perform a Fisher analysis for a range of ΛCDM and dark energy equation of state parameters. We found that when combined with the power spectrum of large-scale structure, the LCF gives rise to strong improvements (up to a factor of two) on the amplitude of fluctuations and bias

parameters over the power spectrum alone. We demonstrated that this is largely due to its independence of linear bias on large scales.

In Chap. 6 we extended the previous analysis by comparing the constraining power of the LCF with the bispectrum and two other three-point statistics: the integrated bispectrum, which is mainly probing very squeezed triangle configurations, and a decomposition of the bispectrum into a set of shapes that together constitute an orthonormal basis. The aim of this study was to see whether any of these alternative three-point statistics can extract a comparable amount of information as the bispectrum, but from a smaller number of configurations as this would lead to a more economic size of the covariance matrix. To that end we measured all four statistical quantities from the same ensemble of simulations as in Chap. 5 and evaluated their predictions using tree-level and one-loop perturbation theory, as well as a halo model approach. Computing the Fisher matrices based on these elements (Fig. 6.9) showed that the inclusion of the bispectrum improves parameter constraints by factors of 1.2 to 1.3 over the power spectrum, and even up to a factor of 5 for parameter combinations that are degenerate, such as σ_8 and b_1. While this analysis neglected shot noise, we found that taking it into account increases the relative importance of three-point statistics (Fig. 6.16). Furthermore, we demonstrated that the modal decomposition gives rise to nearly identical constraints as the generic bispectrum, yet at a fraction of the number of configurations (see Table 6.3) — about 10 modes are sufficient to capture all available three-point information, which reduces the size of the covariance matrix by at least an order of magnitude. In comparison, the LCF gives slightly weaker constraints, while the integrated bispectrum was found to be mostly insensitive to the cosmological parameters used in our analysis. Finally, we studied the impact of using a Gaussian approximation for the covariance matrices and the effect of the various theoretical models on our Fisher forecasts, finding that the former underestimates the parameter uncertainties by nearly a factor of 4, whereas the latter changes the forecasts by 30 % - 50 %.

The final chapter in this thesis (Chap. 7) dealt with the modelling of the galaxy bispectrum in the slightly non-linear regime. In particular, our goal was to extend the current model that has been used in all galaxy survey analyses until today by taking into account all contributions from a complete galaxy bias expansion up to the one-loop level. We first showed how to construct a set of basis operators that encode all effects from galaxy formation on large scales and at each order of perturbation theory. We then illuminated that the problem of bias renormalisation is tightly related to summing over reducible diagrams, which is well known in the context of renormalised perturbation theory. That encouraged us to express the bias expansion in terms of multi-point propagators, which are exactly the renormalised bias parameters. We evaluated the first few multi-point propagators at some initial time, and after evolving them, they serve as the basic building blocks for the power spectrum and bispectrum. We integrated

the resulting bispectrum model into an MCMC likelihood code and fitted it for the first time to measurements from mock galaxy catalogues resembling the *BOSS* CMASS sample. We found that the constraints on the bias parameters from power spectrum and bispectrum are completely consistent, at least up to scales of $k_{\rm max} = 0.2\,h\,{\rm Mpc}^{-1}$, and we showed that the bispectrum model evaluated for the joint best-fit parameters agrees within $\sim 2\,\%$ with the data. Furthermore, we demonstrated that up to the maximal scale and the particular galaxy sample used in our analysis, higher-derivative terms in the bias expansion (which give rise to a scale dependence of the bias parameters) are not required and may even lead to a systematic bias in the results if they are added to the tree-level model without the loop corrections.

8.2 Outlook

Before the bispectrum (or one of the proxies discussed in this thesis) can be robustly applied to the data of future galaxy redshift surveys, further challenges will need to be addressed. Most obviously, the bispectrum model that we presented in Chap. 7 needs to be extended to account for the mapping between real and redshift space. Current analyses of the bispectrum (Gil-Marín et al., 2017) rely on the dispersion model, introduced for the power spectrum in Sec. 3.3.3, which combines the linear Kaiser effect with damping factors due to the velocity dispersion on small scales (Scoccimarro et al., 1999a). However, it was already shown by Scoccimarro (2004) that such models are inherently inaccurate because they do not account for the full non-linearity of the redshift space mapping and thus underestimate the effect of the velocity dispersion on large scales. Improved models for the power spectrum have been developed (e.g. Scoccimarro, 2004; Taruya et al., 2010) and applied to recent data releases of *BOSS* and *eBOSS* (Beutler et al., 2017; Grieb et al., 2017; Hou et al., 2018; Sánchez et al., 2017). An extension of these approaches to the bispectrum has also been shown to be feasible (Hashimoto et al., 2017), and promises to be the way forward.

More accurate theoretical models typically involve a large number of loop integrals and therefore come at the expense of longer computation times. For instance, evaluating all of the bias and matter loop integrals (see Chap. 7) for the galaxy bispectrum at ~ 3000 different triangle configurations takes about 2 - 3 minutes on a single CPU. This becomes problematic once these models are being used to explore the posteriors of a set of cosmological parameters, since each step in parameter space requires the re-computation of all model predictions. Implementing the theory codes into more efficient sampling software that use less likelihood evaluations, combined with interpolation techniques for steps in parameter space that do not stray too far from a given set

of fixed points for which the likelihood was pre-computed, can potentially circumvent this problem. Another promising approach was presented in Simonović et al. (2018), which decomposes the linear power spectrum into a set of power laws and has the advantage that it factorises the cosmology dependence — similar to the bias coefficients in Chap. 7. The remaining loop integrals for each power law only need to be computed once, which would significantly speed up the fitting procedure.

Another problem not discussed in this thesis is the survey selection function. Ignoring shot noise the *measured* bispectrum relates to the *true* one through a convolution with the window function $W(\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3)$, which encodes the angular and radial footprint of the survey:

$$B_{\text{measured}}(k_1, k_2, k_3) = \int \mathrm{d}^3 q_1 \, \mathrm{d}^3 q_2 \, \mathrm{d}^3 q_3 \; W(\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3) \, B_{\text{true}}(\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3) \, \delta_D(\boldsymbol{q}_{123}) \,. \tag{8.1}$$

Hence, when comparing a theoretical model to the measured bispectrum we first have to evaluate the integral in Eq. (8.1). In general this is an expensive operation because the integral cannot be made separable as B_{true} is not separable either. Under certain conditions (tree-level perturbation theory and triangle configurations, which are neither squeezed nor folded), this is still a reasonable approximation as demonstrated by the recent *BOSS* analyses (Gil-Marín et al., 2015, 2017). However, it remains unclear whether the same holds once one-loop corrections have been included, and besides it would be favourable to include as many triangle configurations as possible. In order to account for the k-space binning effect in Chap. 7, we have already developed an interpolation scheme based on Delaunay tesselations that approximates Eq. (8.1) for the particular window function $W_{123} = \tilde{\Pi}_1 \, \tilde{\Pi}_2 \, \tilde{\Pi}_3 / V_B$ (using the notation from Sec. 6.4.1). This interpolation method is almost immediately applicable to any generic window function, which might provide an accurate but still economic enough evaluation of Eq. (8.1).

In Chap. 7 we have seen how a more accurate bias model for the bispectrum inevitably introduces a large number of additional free parameters. In a realistic application these bias parameters need to be marginalised over, which reduces the constraining power on parameters that are cosmologically interesting, such as the growth rate. It is therefore crucial to explore methods that can help reducing uncertainties on the bias parameters, or even fix them a priori. A potential solution are the peak-background split predictions, which can be derived from separate universe simulations as demonstrated in Lazeyras and Schmidt (2017); Lazeyras et al. (2016). In Sec. 7.5.2 we have already compared our measurements for the quadratic bias parameter to such a prediction and found excellent agreement, but it remains to be seen whether fixing these parameters introduces unwanted systematic biases. The combination with the large-scale trispectrum presents another opportunity to significantly reduce uncertainties, especially on

the third-order bias parameters that enter the bispectrum only through loop corrections. As discussed at the end of Sec. 7.4.1, these parameters will appear at tree-level for the trispectrum, which means we can expect improvements comparable to those on b_2 and γ_2 from the bispectrum (see Fig. 7.6).

Clearly, there are many exciting avenues to be pursued during future research, and we hope that this thesis has build a solid foundation for these endeavours.

References

Abell, P. A., Allison, J., Anderson, S. F., et al. Lsst science book, version 2.0. 2009.

Acquaviva, V., Bartolo, N., Matarrese, S., and Riotto, A. Gauge-invariant second-order perturbations and non-Gaussianity from inflation. *Nuclear Physics B*, 667:119–148, September 2003. doi:10.1016/S0550-3213(03)00550-9.

Alam, S., Ata, M., Bailey, S., et al. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample. *MNRAS*, 470:2617–2652, September 2017. doi:10.1093/mnras/stx721.

Albrecht, A., Bernstein, G., Cahn, R., et al. Report of the Dark Energy Task Force. *arXiv:astro-ph/0609591*, September 2006.

Albrecht, A. and Steinhardt, P. J. Cosmology for Grand Unified Theories with Radiatively Induced Symmetry Breaking. *PRL*, 48:1220–1223, April 1982. doi:10.1103/PhysRevLett.48.1220.

Alcock, C. and Paczynski, B. An evolution free test for non-zero cosmological constant. *Nature*, 281:358, October 1979. doi:10.1038/281358a0.

Alpaslan, M., Robotham, A. S. G., Obreschkow, D., et al. Galaxy and Mass Assembly (GAMA): fine filaments of galaxies detected within voids. *MNRAS*, 440:L106–L110, May 2014. doi:10.1093/mnrasl/slu019.

Amendola, L., Appleby, S., Avgoustidis, A., et al. Cosmology and fundamental physics with the Euclid satellite. *Living Reviews in Relativity*, 21:2, April 2018. doi:10.1007/s41114-017-0010-3.

Anderson, T. W. *An Introduction to Multivariate Statistical Analysis*. Wiley, 2003.

Assassi, V., Baumann, D., Green, D., and Zaldarriaga, M. Renormalized halo bias. *Journal of Cosmology and Astro-Particle Physics*, 2014:056, August 2014. doi:10.1088/1475-7516/2014/08/056.

Aubourg, É., Bailey, S., Bautista, J. E., et al. Cosmological implications of baryon acoustic oscillation measurements. *PRD*, 92:123516, December 2015. doi:10.1103/PhysRevD.92.123516.

Baldauf, T., Seljak, U., Desjacques, V., and McDonald, P. Evidence for quadratic tidal tensor bias from the halo bispectrum. *PRD*, 86:083540, October 2012. doi:10.1103/PhysRevD.86.083540.

Baldauf, T., Mercolli, L., Mirbabayi, M., and Pajer, E. The bispectrum in the Effective Field Theory of Large Scale Structure. *Journal of Cosmology and Astro-Particle Physics*, 2015:007, May 2015. doi:10.1088/1475-7516/2015/05/007.

Baldauf, T., Mirbabayi, M., Simonovi, M., and Zaldarriaga, M. LSS constraints with controlled theoretical uncertainties. 2016.

Ballinger, W. E., Peacock, J. A., and Heavens, A. F. Measuring the cosmological constant with redshift surveys. *MNRAS*, 282:877, October 1996. doi:10.1093/mnras/282.3.877.

Bardeen, J. M., Bond, J. R., Kaiser, N., and Szalay, A. S. The Statistics of Peaks of Gaussian Random Fields. *ApJ*, 304:15, May 1986. doi:10.1086/164143.

Bardeen, J. M., Steinhardt, P. J., and Turner, M. S. Spontaneous creation of almost scale-free density perturbations in an inflationary universe. *PRD*, 28:679–693, August 1983. doi:10.1103/PhysRevD.28.679.

Baumann, D. TASI Lectures on Inflation. *ArXiv e-prints*, art. arXiv:0907.5424, July 2009.

Baumann, D., Nicolis, A., Senatore, L., and Zaldarriaga, M. Cosmological non-linearities as an effective fluid. *Journal of Cosmology and Astro-Particle Physics*, 2012:051, July 2012. doi:10.1088/1475-7516/2012/07/051.

Baumgart, D. J. and Fry, J. N. Fourier Spectra of Three-dimensional Data. *ApJ*, 375:25, July 1991. doi:10.1086/170166.

Bel, J., Hoffmann, K., and Gaztañaga, E. Non-local bias contribution to third-order galaxy correlations. *MNRAS*, 453:259–276, October 2015. doi:10.1093/mnras/stv1600.

Bennett, C. L., Banday, A. J., Gorski, K. M., et al. Four-Year COBE DMR Cosmic Microwave Background Observations: Maps and Basic Results. *ApJ*, 464:L1, June 1996. doi:10.1086/310075.

Bennett, C. L., Larson, D., Weiland, J. L., et al. Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results. *The Astrophysical Journal Supplement Series*, 208:20, October 2013. doi:10.1088/0067-0049/208/2/20.

Berlind, A. A. and Weinberg, D. H. The Halo Occupation Distribution: Toward an Empirical Determination of the Relation between Galaxies and Mass. *ApJ*, 575:587–616, August 2002. doi:10.1086/341469.

Bernardeau, F., Colombi, S., Gaztañaga, E., and Scoccimarro, R. Large-scale structure of the Universe and cosmological perturbation theory. *Phys. Rep.* , 367:1–248, September 2002. doi:10.1016/S0370-1573(02)00135-7.

Bernardeau, F., Crocce, M., and Sefusatti, E. Multipoint propagators for non-Gaussian initial conditions. *PRD*, 82(8):083507, October 2010. doi:10.1103/PhysRevD.82.083507.

Bernardeau, F., Crocce, M., and Scoccimarro, R. Multipoint propagators in cosmological gravitational instability. *PRD*, 78:103521, November 2008. doi:10.1103/PhysRevD.78.103521.

Bernardeau, F., Crocce, M., and Scoccimarro, R. Constructing regularized cosmic propagators. *PRD*, 85:123519, June 2012. doi:10.1103/PhysRevD.85.123519.

Beutler, F., Blake, C., Colless, M., et al. The 6dF Galaxy Survey: baryon acoustic oscillations and the local Hubble constant. *MNRAS*, 416:3017–3032, October 2011. doi:10.1111/j.1365-2966.2011.19250.x.

Beutler, F., Seo, H.-J., Saito, S., et al. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: anisotropic galaxy clustering in Fourier space. *MNRAS*, 466:2242–2260, April 2017. doi:10.1093/mnras/stw3298.

Blake, C. and Glazebrook, K. Probing Dark Energy Using Baryonic Oscillations in the Galaxy Power Spectrum as a Cosmological Ruler. *ApJ*, 594:665–673, September 2003. doi:10.1086/376983.

Blot, L., Corasaniti, P. S., Amendola, L., and Kitching, T. D. Non-Linear Matter Power Spectrum Covariance Matrix Errors and Cosmological Parameter Uncertainties. *Mon. Not. Roy. Astron. Soc.*, 458(4):4462–4470, 2016. doi:10.1093/mnras/stw604.

Bond, J. R., Cole, S., Efstathiou, G., and Kaiser, N. Excursion Set Mass Functions for Hierarchical Gaussian Fluctuations. *ApJ*, 379:440, October 1991. doi:10.1086/170520.

Bouchet, F. R., Juszkiewicz, R., Colombi, S., and Pellat, R. Weakly nonlinear gravitational instability for arbitrary Omega. *ApJL*, 394:L5–L8, July 1992. doi:10.1086/186459.

Bouchet, F. R., Colombi, S., Hivon, E., and Juszkiewicz, R. Perturbative Lagrangian approach to gravitational instability. *A&A*, 296:575, April 1995.

Boylan-Kolchin, M., Bullock, J. S., and Kaplinghat, M. The Milky Way's bright satellites as an apparent failure of ΛCDM. *MNRAS*, 422:1203–1218, May 2012. doi:10.1111/j.1365-2966.2012.20695.x.

Buchert, T. Lagrangian Theory of Gravitational Instability of Friedman-Lemaitre Cosmologies - a Generic Third-Order Model for Nonlinear Clustering. *MNRAS*, 267:811, April 1994. doi:10.1093/mnras/267.4.811.

Buchert, T., Melott, A. L., and Weiss, A. G. Testing higher-order Lagrangian perturbation theory against numerical simulations I. Pancake models. *A&A*, 288:349–364, August 1994.

Bull, P., Akrami, Y., Adamek, J., et al. Beyond Λ CDM: Problems, solutions, and the road ahead. *Physics of the Dark Universe*, 12:56–99, June 2016. doi:10.1016/j.dark.2016.02.001.

Burgess, C. P. The Cosmological Constant Problem: Why it's hard to get Dark Energy from Micro-physics. *ArXiv e-prints*, art. arXiv:1309.4133, September 2013.

Byun, J., Eggemeier, A., Regan, D., Seery, D., and Smith, R. E. Towards optimal cosmological parameter recovery from compressed bispectrum statistics. *MNRAS*, 471: 1581–1618, October 2017. doi:10.1093/mnras/stx1681.

Carlson, J., White, M., and Padmanabhan, N. Critical look at cosmological perturbation theory techniques. *PRD*, 80:043531, August 2009. doi:10.1103/PhysRevD.80.043531.

Carrasco, J. J. M., Hertzberg, M. P., and Senatore, L. The effective field theory of cosmological large scale structures. *Journal of High Energy Physics*, 2012:82, September 2012. doi:10.1007/JHEP09(2012)082.

Carron, J. and Szapudi, I. Optimal non-linear transformations for large scale structure statistics. *Mon. Not. Roy. Astron. Soc.*, 434:2961, 2013. doi:10.1093/mnras/stt1215.

Catelan, P., Porciani, C., and Kamionkowski, M. Two ways of biasing galaxy formation. *MNRAS*, 318:L39–L44, November 2000. doi:10.1046/j.1365-8711.2000.04023.x.

Catelan, P. Lagrangian dynamics in non-flat universes and non-linear gravitational evolution. *MNRAS*, 276:115–124, September 1995. doi:10.1093/mnras/276.1.115.

Chan, K. C., Scoccimarro, R., and Sheth, R. K. Gravity and large-scale nonlocal bias. *PRD*, 85(8):083509, April 2012. doi:10.1103/PhysRevD.85.083509.

Chan, K. C. and Blot, L. Assessment of the information content of the power spectrum and bispectrum. *PRD*, 96:023528, July 2017. doi:10.1103/PhysRevD.96.023528.

Chevallier, M. and Polarski, D. Accelerating universes with scaling dark matter. *Int. J. Mod. Phys.*, D10:213–224, 2001. doi:10.1142/S0218271801000822.

Chiang, C.-T., Wagner, C., Schmidt, F., and Komatsu, E. Position-dependent power spectrum of the large-scale structure: a novel method to measure the squeezed-limit bispectrum. *Journal of Cosmology and Astro-Particle Physics*, 5:048, May 2014. doi:10.1088/1475-7516/2014/05/048.

Chiang, C.-T. *Position-dependent power spectrum: a new observable in the large-scale structure*. PhD thesis, Munich U., 2015. URL https://inspirehep.net/record/1387739/files/arXiv:1508.03256.pdf.

Chiang, C.-T., Wagner, C., Sánchez, A. G., Schmidt, F., and Komatsu, E. Position-dependent correlation function from the SDSS-III Baryon Oscillation Spectroscopic Survey Data Release 10 CMASS Sample. *JCAP*, 1509(09):028, 2015. doi:10.1088/1475-7516/2015/09/028, 10.1088/1475-7516/2015/9/028.

Chiang, L.-Y. The importance of Fourier phases for the morphology of gravitational clustering. *MNRAS*, 325:405–411, July 2001. doi:10.1046/j.1365-8711.2001.04458.x.

Chiang, L.-Y. and Coles, P. Phase information and the evolution of cosmological density perturbations. *MNRAS*, 311:809–824, February 2000. doi:10.1046/j.1365-8711.2000.03086.x.

Clifton, T., Ferreira, P. G., Padilla, A., and Skordis, C. Modified gravity and cosmology. *Phys. Rep.* , 513:1–189, March 2012. doi:10.1016/j.physrep.2012.01.001.

Cole, S., Percival, W. J., Peacock, J. A., et al. The 2dF Galaxy Redshift Survey: power-spectrum analysis of the final data set and cosmological implications. *MNRAS*, 362: 505–534, September 2005. doi:10.1111/j.1365-2966.2005.09318.x.

Coles, P. Galaxy formation with a local bias. *MNRAS*, 262:1065–1075, June 1993. doi:10.1093/mnras/262.4.1065.

Coles, P. and Chiang, L.-Y. Characterizing the nonlinear growth of large-scale structure in the Universe. *Nature*, 406:376–378, July 2000. doi:10.1038/35019009.

Colless, M., Dalton, G., Maddox, S., et al. The 2dF Galaxy Redshift Survey: spectra and redshifts. *MNRAS*, 328:1039–1063, December 2001. doi:10.1046/j.1365-8711.2001.04902.x.

Colless, M., Peterson, B. A., Jackson, C., et al. The 2dF Galaxy Redshift Survey: Final Data Release. *ArXiv e-prints*, art. astro-ph/0306581, June 2003.

Cooray, A. and Sheth, R. Halo models of large scale structure. *Phys. Rep.*, 372:1–129, December 2002.

Crocce, M. and Scoccimarro, R. Memory of initial conditions in gravitational clustering. *PRD*, 73(6):063520, March 2006a. doi:10.1103/PhysRevD.73.063520.

Crocce, M., Pueblas, S., and Scoccimarro, R. Transients from Initial Conditions in Cosmological Simulations. *Mon. Not. Roy. Astron. Soc.*, 373:369–381, 2006. doi:10.1111/j.1365-2966.2006.11040.x.

Crocce, M., Scoccimarro, R., and Bernardeau, F. MPTBREEZE: a fast renormalized perturbative scheme. *MNRAS*, 427:2537–2551, December 2012. doi:10.1111/j.1365-2966.2012.22127.x.

Crocce, M. and Scoccimarro, R. Renormalized cosmological perturbation theory. *PRD*, 73:063519, March 2006b. doi:10.1103/PhysRevD.73.063519.

Crocce, M. and Scoccimarro, R. Nonlinear evolution of baryon acoustic oscillations. *PRD*, 77:023533, January 2008. doi:10.1103/PhysRevD.77.023533.

D'Amico, G. and Scoccimarro, R. *in prep.*, 2018.

Davis, M. and Peebles, P. J. E. A survey of galaxy redshifts. V. The two-point position and velocity correlations. *ApJ*, 267:465–482, April 1983. doi:10.1086/160884.

Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. The Baryon Oscillation Spectroscopic Survey of SDSS-III. AJ, 145:10, January 2013. doi:10.1088/0004-6256/145/1/10.

de Jong, R. S., Barden, S. C., Bellido-Tirado, O., et al. 4MOST: the 4-metre Multi-Object Spectroscopic Telescope project at preliminary design review. In *Ground-based and Airborne Instrumentation for Astronomy VI*, volume 9908, page 99081O, August 2016. doi:10.1117/12.2232832.

DES Collaboration, Abbott, T. M. C., Abdalla, F. B., et al. Dark Energy Survey Year 1 Results: Cosmological Constraints from Galaxy Clustering and Weak Lensing. *ArXiv e-prints*, art. arXiv:1708.01530, August 2017.

DESI Collaboration, Aghamousa, A., Aguilar, J., et al. The DESI Experiment Part I: Science,Targeting, and Survey Design. *ArXiv e-prints*, art. arXiv:1611.00036, October 2016.

Desjacques, V., Jeong, D., and Schmidt, F. Large-scale galaxy bias. *Phys. Rep.* , 733:1–193, February 2018. doi:10.1016/j.physrep.2017.12.002.

Devroye, L. *Non-Uniform Random Variate Generation*. Springer, 1986.

Dewdney, P. E., Hall, P. J., Schilizzi, R. T., and Lazio, T. J. L. W. The Square Kilometre Array. *IEEE Proceedings*, 97:1482–1496, August 2009. doi:10.1109/JPROC.2009.2021005.

Dodelson, S. *Modern Cosmology*. Academic Press, 2011.

Driver, S. P., Norberg, P., Baldry, I. K., et al. GAMA: towards a physical understanding of galaxy formation. *Astronomy and Geophysics*, 50:5.12–5.19, October 2009. doi:10.1111/j.1468-4004.2009.50512.x.

Dubinski, J. and Carlberg, R. G. The Structure of Cold Dark Matter Halos. *ApJ*, 378:496, September 1991. doi:10.1086/170451.

Efstathiou, G. Constraining the equation of state of the universe from distant type Ia supernovae and cosmic microwave background anisotropies. *Mon. Not. Roy. Astron. Soc.*, 310:842–850, 1999. doi:10.1046/j.1365-8711.1999.02997.x.

Eggemeier, A., Battefeld, T., Smith, R. E., and Niemeyer, J. The anisotropic line correlation function as a probe of anisotropies in galaxy surveys. *MNRAS*, 453:797–809, October 2015. doi:10.1093/mnras/stv1602.

Eggemeier, A. and Smith, R. E. Cosmology with Phase Statistics: Parameter Forecasts and Detectability of BAO. *Mon. Not. Roy. Astron. Soc.*, 466:2496, 2017. doi:10.1093/mnras/stw3249.

Einasto, J., Kaasik, A., and Saar, E. Dynamic evidence on massive coronas of galaxies. *Nature*, 250:309–310, July 1974. doi:10.1038/250309a0.

Eisenstein, D. J. and Hu, W. Baryonic Features in the Matter Transfer Function. *ApJ*, 496: 605–614, March 1998. doi:10.1086/305424.

Eisenstein, D. J., Hu, W., and Tegmark, M. Cosmic Complementarity: H_0 and Ω_m from Combining Cosmic Microwave Background Experiments and Redshift Surveys. *ApJ*, 504:L57–L60, September 1998. doi:10.1086/311582.

Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies. *ApJ*, 633: 560–574, November 2005. doi:10.1086/466512.

Feldman, H. A., Kaiser, N., and Peacock, J. A. Power-Spectrum Analysis of Three-dimensional Redshift Surveys. *ApJ*, 426:23, May 1994. doi:10.1086/174036.

Fergusson, J. R. and Shellard, E. P. S. Shape of primordial non-Gaussianity and the CMB bispectrum. *Phys. Rev.*, D80(4):043510, August 2009. doi:10.1103/PhysRevD.80.043510.

Fergusson, J. R., Liguori, M., and Shellard, E. P. S. General CMB and Primordial Bispectrum Estimation I: Mode Expansion, Map-Making and Measures of $f_N L$. *Phys. Rev.*, D82:023502, 2010. doi:10.1103/PhysRevD.82.023502.

Fergusson, J. R., Regan, D. M., and Shellard, E. P. S. Rapid Separable Analysis of Higher Order Correlators in Large Scale Structure. *Phys. Rev.*, D86:063511, 2012. doi:10.1103/PhysRevD.86.063511.

Fixsen, D. J. The Temperature of the Cosmic Microwave Background. *ApJ*, 707:916–920, December 2009. doi:10.1088/0004-637X/707/2/916.

Franco, F. O., Bonvin, C., Obreschkow, D., and Ali, K. Probing redshift-space distortions with phase correlations. *ArXiv e-prints*, art. arXiv:1805.10178, May 2018.

Friedmann, A. Über die Krümmung des Raumes. *Zeitschrift fur Physik*, 10:377–386, January 1922. doi:10.1007/BF01332580.

Frieman, J. A. and Gaztañaga, E. The Projected Three-Point Correlation Function: Theory and Observations. *ApJ*, 521:L83–L86, August 1999. doi:10.1086/312195.

Frieman, J. A. and Gaztanaga, E. The Three-Point Function as a Probe of Models for Large-Scale Structure. *ApJ*, 425:392, April 1994. doi:10.1086/173995.

Fry, J. N. The Galaxy correlation hierarchy in perturbation theory. *ApJ*, 279:499–510, April 1984. doi:10.1086/161913.

Fry, J. N. Gravity, bias, and the galaxy three-point correlation function. *PRL*, 73:215–219, July 1994. doi:10.1103/PhysRevLett.73.215.

Fry, J. N. The Evolution of Bias. *ApJ*, 461:L65, April 1996. doi:10.1086/310006.

Fry, J. N. and Gaztanaga, E. Biasing and hierarchical statistics in large-scale structure. *ApJ*, 413:447–452, August 1993. doi:10.1086/173015.

Fry, J. N. and Seldner, M. Transform analysis of the high-resolution Shane-Wirtanen Catalog - The power spectrum and the bispectrum. *ApJ*, 259:474–481, August 1982. doi:10.1086/160184.

Fry, J. N., Melott, A. L., and Shandarin, S. F. The Three-Point Correlation Function in an Ensemble of Three-dimensional Simulations. *ApJ*, 412:504, August 1993. doi:10.1086/172938.

Gagrani, P. and Samushia, L. Information Content of the Angular Multipoles of Redshift-Space Galaxy Bispectrum. *MNRAS*, 467:928–935, May 2017. doi:10.1093/mnras/stx135.

Gaztanaga, E. and Frieman, J. A. Bias and high-order galaxy correlation functions in the APM galaxy survey. *ApJL*, 437:L13–L16, December 1994. doi:10.1086/187671.

Gil-Marín, H., Noreña, J., Verde, L., et al. The power spectrum and bispectrum of SDSS DR11 BOSS galaxies - I. Bias and gravity. *MNRAS*, 451:539–580, July 2015. doi:10.1093/mnras/stv961.

Gil-Marín, H., Percival, W. J., Verde, L., et al. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: RSD measurement from the power spectrum and bispectrum of the DR12 BOSS galaxies. *MNRAS*, 465:1757–1788, February 2017. doi:10.1093/mnras/stw2679.

Glazebrook, K. and Blake, C. Measuring the Cosmic Evolution of Dark Energy with Baryonic Oscillations in the Galaxy Power Spectrum. *ApJ*, 631:1–20, September 2005. doi:10.1086/432497.

Goroff, M. H., Grinstein, B., Rey, S. J., and Wise, M. B. Coupling of modes of cosmological mass density fluctuations. *ApJ*, 311:6–14, December 1986. doi:10.1086/164749.

Gott, I., J. Richard, Melott, A. L., and Dickinson, M. The Sponge-like Topology of Large-Scale Structure in the Universe. *ApJ*, 306:341, July 1986. doi:10.1086/164347.

Grieb, J. N., Sánchez, A. G., Salazar-Albornoz, S., and Dalla Vecchia, C. Gaussian covariance matrices for anisotropic galaxy clustering measurements. *MNRAS*, 457: 1577–1592, April 2016. doi:10.1093/mnras/stw065.

Grieb, J. N., Sánchez, A. G., Salazar-Albornoz, S., et al. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: Cosmological implications of the Fourier space wedges of the final sample. *MNRAS*, 467:2085–2112, May 2017. doi:10.1093/mnras/stw3384.

Groth, E. J. and Peebles, P. J. E. Statistical analysis of catalogs of extragalactic objects. VII. Two- and three-point correlation functions for the high-resolution Shane- Wirtanen catalog of galaxies. *ApJ*, 217:385–405, October 1977. doi:10.1086/155588.

Gunn, J. E. and Tinsley, B. M. An accelerating Universe. *Nature*, 257:454–457, October 1975. doi:10.1038/257454a0.

Guo, H., Zheng, Z., Zehavi, I., et al. Redshift-space clustering of SDSS galaxies - luminosity dependence, halo occupation distribution, and velocity bias. *MNRAS*, 453: 4368–4383, November 2015. doi:10.1093/mnras/stv1966.

Guth, A. H. Inflationary universe: A possible solution to the horizon and flatness problems. *PRD*, 23:347–356, January 1981. doi:10.1103/PhysRevD.23.347.

Guth, A. H. and Pi, S.-Y. Fluctuations in the New Inflationary Universe. *PRL*, 49:1110–1113, October 1982. doi:10.1103/PhysRevLett.49.1110.

Guth, A. H. and Pi, S.-Y. Quantum mechanics of the scalar field in the new inflationary universe. *PRD*, 32:1899–1920, October 1985. doi:10.1103/PhysRevD.32.1899.

Hahn, C., Beutler, F., Sinha, M., et al. Likelihood Non-Gaussianity in Large-Scale Structure Analyses. *ArXiv e-prints*, art. arXiv:1803.06348, March 2018.

Hahn, T. Concurrent Cuba. *Comput. Phys. Commun.*, 207:341–349, 2016. doi:10.1016/j.cpc.2016.05.012.

Hamilton, A. J. S. Measuring Omega and the Real Correlation Function from the Redshift Correlation Function. *ApJ*, 385:L5, January 1992. doi:10.1086/186264.

Hamilton, A. J. S. Linear Redshift Distortions: a Review. In *The Evolving Universe*, volume 231, page 185, January 1998. doi:10.1007/978-94-011-4960-0_17.

Hamilton, A. J. S. Formulae for growth factors in expanding universes containing matter and a cosmological constant. *MNRAS*, 322:419–425, April 2001. doi:10.1046/j.1365-8711.2001.04137.x.

Handley, W. J., Hobson, M. P., and Lasenby, A. N. POLYCHORD: next-generation nested sampling. *MNRAS*, 453:4384–4398, November 2015. doi:10.1093/mnras/stv1911.

Hartlap, J., Simon, P., and Schneider, P. Why your model parameter confidences might be too optimistic. Unbiased estimation of the inverse covariance matrix. *A&A*, 464:399–404, March 2007. doi:10.1051/0004-6361:20066170.

Hashimoto, I., Rasera, Y., and Taruya, A. Precision cosmology with redshift-space bispectrum: A perturbation theory based model at one-loop order. *PRD*, 96:043526, August 2017. doi:10.1103/PhysRevD.96.043526.

Hauser, M. G. and Peebles, P. J. E. Statistical Analysis of Catalogs of Extragalactic Objects. II. the Abell Catalog of Rich Clusters. *ApJ*, 185:757–786, November 1973. doi:10.1086/152453.

Hawking, S. W. The development of irregularities in a single bubble inflationary universe. *Physics Letters B*, 115:295–297, September 1982. doi:10.1016/0370-2693(82)90373-2.

Hawkins, E., Maddox, S., Cole, S., et al. The 2dF Galaxy Redshift Survey: correlation functions, peculiar velocities and the matter density of the Universe. *MNRAS*, 346:78–96, November 2003. doi:10.1046/j.1365-2966.2003.07063.x.

Hearin, A. P., Zentner, A. R., van den Bosch, F. C., Campbell, D., and Tollerud, E. Introducing decorated HODs: modelling assembly bias in the galaxy-halo connection. *MNRAS*, 460:2552–2570, August 2016. doi:10.1093/mnras/stw840.

Hikage, C., Matsubara, T., and Suto, Y. The Distribution Function of the Phase Sum as a Signature of Phase Correlations Induced by Nonlinear Gravitational Clustering. *ApJ*, 600:553–563, January 2004. doi:10.1086/379851.

Hikage, C., Suto, Y., Kayo, I., et al. Three-Dimensional Genus Statistics of Galaxies in the SDSS Early Data Release. *Publications of the Astronomical Society of Japan*, 54:707–717, October 2002. doi:10.1093/pasj/54.5.707.

Hikage, C., Schmalzing, J., Buchert, T., et al. Minkowski Functionals of SDSS Galaxies I : Analysis of Excursion Sets. *Publications of the Astronomical Society of Japan*, 55: 911–931, October 2003. doi:10.1093/pasj/55.5.911.

Hikage, C., Matsubara, T., Suto, Y., et al. Fourier Phase Analysis of SDSS Galaxies. *Publications of the Astronomical Society of Japan*, 57:709–718, October 2005. doi:10.1093/pasj/57.5.709.

Hivon, E., Bouchet, F. R., Colombi, S., and Juszkiewicz, R. Redshift distortions of clustering: a Lagrangian approach. *A&A*, 298:643, June 1995.

Hoffmann, K., Bel, J., and Gaztañaga, E. Comparing halo bias from abundance and clustering. *MNRAS*, 450:1674–1692, June 2015. doi:10.1093/mnras/stv702.

Hou, J., Sánchez, A. G., Scoccimarro, R., et al. The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample: anisotropic clustering analysis in configuration-space. *ArXiv e-prints*, art. arXiv:1801.02656, January 2018.

Hoyle, F., Vogeley, M. S., Gott, I., J. Richard, et al. Two-dimensional Topology of the Sloan Digital Sky Survey. *ApJ*, 580:663–671, December 2002. doi:10.1086/343734.

Hu, W. Lecture notes on Cosmology.

Hubble, E. Extragalactic nebulae. *ApJ*, 64:321–369, December 1926. doi:10.1086/143018.

Hubble, E. A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae. *Proceedings of the National Academy of Science*, 15:168–173, March 1929. doi:10.1073/pnas.15.3.168.

Hubble, E. The Distribution of Extra-Galactic Nebulae. *ApJ*, 79:8, January 1934. doi:10.1086/143517.

Huchra, J., Davis, M., Latham, D., and Tonry, J. X-Ray Spectra of Active Galactic Nuclei. *The Astrophysical Journal Supplement Series*, 52:89, June 1983. doi:10.1086/190860.

Huterer, D. and Shafer, D. L. Dark energy two decades after: observables, probes, consistency tests. *Reports on Progress in Physics*, 81:016901, January 2018. doi:10.1088/1361-6633/aa997e.

Jackson, J. C. A critique of Rees's theory of primordial gravitational radiation. *MNRAS*, 156:1P, January 1972. doi:10.1093/mnras/156.1.1P.

Jain, B. and Bertschinger, E. Second-Order Power Spectrum and Nonlinear Evolution at High Redshift. *ApJ*, 431:495, August 1994. doi:10.1086/174502.

Jain, B. and Bertschinger, E. Self-Similar Evolution of Gravitational Clustering. II. N-Body Simulations of the N = -2 Spectrum. *ApJ*, 509:517–530, December 1998. doi:10.1086/306538.

Jenkins, A., Frenk, C. S., White, S. D. M., et al. The mass function of dark matter haloes. *MNRAS*, 321:372–384, February 2001. doi:10.1046/j.1365-8711.2001.04029.x.

Jeong, D. and Komatsu, E. Perturbation Theory Reloaded: Analytical Calculation of Nonlinearity in Baryonic Oscillations in the Real-Space Matter Power Spectrum. *ApJ*, 651:619–626, November 2006. doi:10.1086/507781.

Joachimi, B., Shi, X., and Schneider, P. Bispectrum covariance in the flat-sky limit. *Astron. Astrophys.*, 508:1193–1204, December 2009. doi:10.1051/0004-6361/200912906.

Juszkiewicz, R. On the evolution of cosmological adiabatic perturbations in the weakly non-linear regime. *MNRAS*, 197:931–940, December 1981. doi:10.1093/mnras/197.4.931.

Kaiser, N. On the spatial correlations of Abell clusters. *ApJ*, 284:L9–L12, September 1984. doi:10.1086/184341.

Kaiser, N. Clustering in real space and in redshift space. *MNRAS*, 227:1–21, July 1987. doi:10.1093/mnras/227.1.1.

Katz, L. and Mulders, G. F. W. On the Clustering of Nebulae. II. *ApJ*, 95:565, May 1942. doi:10.1086/144423.

Kayo, I., Takada, M., and Jain, B. Information content of weak lensing power spectrum and bispectrum: including the non-Gaussian error covariance matrix. *Mon. Not. Roy. Astron. Soc.*, 429:344–371, 2013. doi:10.1093/mnras/sts340.

Klypin, A. and Prada, F. Dark matter statistics for large galaxy catalogs: power spectra and covariance matrices. *MNRAS*, page 1328, June 2018. doi:10.1093/mnras/sty1340.

Klypin, A., Kravtsov, A. V., Valenzuela, O., and Prada, F. Where Are the Missing Galactic Satellites? *ApJ*, 522:82–92, September 1999. doi:10.1086/307643.

Kofman, L. and Pogosyan, D. Dynamics of gravitational instability is nonlocal. *ApJ*, 442: 30–38, March 1995. doi:10.1086/175419.

Koyama, K. Cosmological tests of modified gravity. *Reports on Progress in Physics*, 79: 046902, April 2016. doi:10.1088/0034-4885/79/4/046902.

Kravtsov, A. V., Berlind, A. A., Wechsler, R. H., et al. The Dark Side of the Halo Occupation Distribution. *ApJ*, 609:35–49, July 2004. doi:10.1086/420959.

Lazanu, A., Giannantonio, T., Schmittfull, M., and Shellard, E. P. S. Matter bispectrum of large-scale structure: Three-dimensional comparison between theoretical models and numerical simulations. *PRD*, 93:083517, April 2016. doi:10.1103/PhysRevD.93.083517.

Lazeyras, T. and Schmidt, F. Beyond LIMD bias: a measurement of the complete set of third-order halo bias parameters. *ArXiv e-prints*, art. arXiv:1712.07531, December 2017.

Lazeyras, T., Wagner, C., Baldauf, T., and Schmidt, F. Precision measurement of the local bias of dark matter halos. *Journal of Cosmology and Astro-Particle Physics*, 2016:018, February 2016. doi:10.1088/1475-7516/2016/02/018.

Lemaître, G. Expansion of the universe, The expanding universe. *MNRAS*, 91:490–501, March 1931a. doi:10.1093/mnras/91.5.490.

Lemaître, G. Expansion of the universe, A homogeneous universe of constant mass and increasing radius accounting for the radial velocity of extra- galactic nebulae. *MNRAS*, 91:483–490, March 1931b. doi:10.1093/mnras/91.5.483.

Lemaitre, G. Evolution of the Expanding Universe. *Proceedings of the National Academy of Science*, 20:12–17, January 1934. doi:10.1073/pnas.20.1.12.

Lewis, A., Challinor, A., and Lasenby, A. Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models. *ApJ*, 538: 473–476, August 2000. doi:10.1086/309179.

Limber, D. N. The Analysis of Counts of the Extragalactic Nebulae in Terms of a Fluctuating Density Field. II. *ApJ*, 119:655, May 1954. doi:10.1086/145870.

Linde, A. D. A new inflationary universe scenario: A possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems. *Physics Letters B*, 108:389–393, February 1982. doi:10.1016/0370-2693(82)91219-9.

Lokas, E. L., Juszkiewicz, R., Bouchet, F. R., and Hivon, E. Previrialization: perturbative and n-body results. *Astrophys. J.*, 467:1, 1996. doi:10.1086/177580.

Macaulay, E., Wehus, I. K., and Eriksen, H. K. Lower Growth Rate from Recent Redshift Space Distortion Measurements than Expected from Planck. *PRL*, 111:161301, October 2013. doi:10.1103/PhysRevLett.111.161301.

Maddox, S. J., Sutherland, W. J., Efstathiou, G., and Loveday, J. The APM galaxy survey - I. APM measurements and star-galaxy separation. *MNRAS*, 243:692–712, April 1990.

Makino, N., Sasaki, M., and Suto, Y. Analytic approach to the perturbative expansion of nonlinear gravitational fluctuations in cosmological density and velocity fields. *Phys. Rev.*, D46:585–602, 1992. doi:10.1103/PhysRevD.46.585.

Maldacena, J. Non-gaussian features of primordial fluctuations in single field inflationary models. *Journal of High Energy Physics*, 2003, May 2003. doi:10.1088/1126-6708/2003/05/013.

Manera, M., Sheth, R. K., and Scoccimarro, R. Large-scale bias and the inaccuracy of the peak-background split. *MNRAS*, 402:589–602, February 2010. doi:10.1111/j.1365-2966.2009.15921.x.

Mardia, K. V. and Jupp, P. E. *Directional Statistics*. Wiley, 1st edition, 1999.

Marin, F. A. et al. The WiggleZ Dark Energy Survey: constraining galaxy bias and cosmic growth with 3-point correlation functions. *Mon. Not. Roy. Astron. Soc.*, 432:2654, 2013. doi:10.1093/mnras/stt520.

Matarrese, S., Verde, L., and Heavens, A. F. Large-scale bias in the Universe: bispectrum method. *MNRAS*, 290:651–662, October 1997. doi:10.1093/mnras/290.4.651.

Matsubara, T. Diagrammatic Methods in Statistics and Biasing in the Large-Scale Structure of the Universe. *ApJS*, 101:1, November 1995. doi:10.1086/192231.

Matsubara, T. Phase correlations in non-Gaussian fields. *Astrophys.J.*, 591:L79–L82, 2003. doi:10.1086/377293.

Matsubara, T. Statistics of Fourier Modes in Non-Gaussian Fields. *The Astrophysical Journal Supplement Series*, 170:1–32, May 2007. doi:10.1086/513466.

Matsubara, T. and Suto, Y. Cosmological Redshift Distortion of Correlation Functions as a Probe of the Density Parameter and the Cosmological Constant. *ApJ*, 470:L1, October 1996. doi:10.1086/310290.

McDonald, P. Clustering of dark matter tracers: Renormalizing the bias parameters. *PRD*, 74(10):103512, November 2006. doi:10.1103/PhysRevD.74.103512.

McDonald, P. and Roy, A. Clustering of dark matter tracers: generalizing bias for the coming era of precision LSS. *Journal of Cosmology and Astro-Particle Physics*, 8:020, August 2009. doi:10.1088/1475-7516/2009/08/020.

Mead, A., Peacock, J., Heymans, C., Joudaki, S., and Heavens, A. An accurate halo model for fitting non-linear cosmological power spectra and baryonic feedback models. *Mon. Not. Roy. Astron. Soc.*, 454(2):1958–1975, 2015. doi:10.1093/mnras/stv2036.

Mead, A., Heymans, C., Lombriser, L., et al. Accurate halo-model matter power spectra with dark energy, massive neutrinos and modified gravitational forces. *Mon. Not. Roy. Astron. Soc.*, 459(2):1468–1488, 2016. doi:10.1093/mnras/stw681.

Mecke, K. R., Buchert, T., and Wagner, H. Robust morphological measures for large-scale structure in the Universe. *A&A*, 288:697–704, August 1994.

Meiksin, A. and White, M. The growth of correlations in the matter power spectrum. *MNRAS*, 308:1179–1184, October 1999. doi:10.1046/j.1365-8711.1999.02825.x.

Mészáros, P. The behaviour of point masses in an expanding cosmological substratum. *A&A*, 37:225–228, December 1974.

Milne, E. A. World-Structure and the Expansion of the Universe. Mit 6 Abbildungen. *Zeitschrift fur Astrophysik*, 6:1, January 1933.

Milne, E. A. A Newtonian expanding Universe. *The Quarterly Journal of Mathematics*, 5: 64–72, January 1934. doi:10.1093/qmath/os-5.1.64.

Mirbabayi, M., Schmidt, F., and Zaldarriaga, M. Biased tracers and time evolution. *Journal of Cosmology and Astro-Particle Physics*, 7:030, July 2015. doi:10.1088/1475-7516/2015/07/030.

Mo, H., van den Bosch, F. C., and White, S. *Galaxy Formation and Evolution*. May 2010.

Mo, H. J. and White, S. D. M. An analytic model for the spatial clustering of dark matter haloes. *MNRAS*, 282:347–361, September 1996. doi:10.1093/mnras/282.2.347.

Monaco, P. Approximate Methods for the Generation of Dark Matter Halo Catalogs in the Age of Precision Cosmology. *Galaxies*, 4:53, October 2016. doi:10.3390/galaxies4040053.

Moore, B., Ghigna, S., Governato, F., et al. Dark Matter Substructure within Galactic Halos. *ApJ*, 524:L19–L22, October 1999. doi:10.1086/312287.

Mortonson, M. J., Weinberg, D. H., and White, M. Dark Energy: A Short Review. *ArXiv e-prints*, art. arXiv:1401.0046, December 2013.

Munshi, D. and Heavens, A. A new approach to probing primordial non-Gaussianity. *MNRAS*, 401:2406–2418, February 2010. doi:10.1111/j.1365-2966.2009.15820.x.

Navarro, J. F., Frenk, C. S., and White, S. D. M. The Structure of cold dark matter halos. *Astrophys. J.*, 462:563–575, 1996. doi:10.1086/177173.

Nernst, W. Über einen Versuch von quantentheoretischen Betrachtungen zur Annahme stetiger Energieänderungen zurückzukehren. *Verhandlungen der Deutschen Physikalischen Gesellschaft*, 18:83, 1916.

Neyman, J. and Scott, E. L. A Theory of the Spatial Distribution of Galaxies. *ApJ*, 116:144, July 1952. doi:10.1086/145599.

Neyrinck, M. C., Szapudi, I., and Szalay, A. S. Rejuvenating the matter power spectrum: restoring information with a logarithmic density mapping. *Astrophys. J.*, 698:L90–L93, 2009. doi:10.1088/0004-637X/698/2/L90.

Norberg, P., Baugh, C. M., Gaztañaga, E., and Croton, D. J. Statistical analysis of galaxy surveys - I. Robust error estimation for two-point clustering statistics. *MNRAS*, 396: 19–38, June 2009. doi:10.1111/j.1365-2966.2009.14389.x.

Obreschkow, D., Power, C., Bruderer, M., and Bonvin, C. A Robust Measure of Cosmic Structure beyond the Power Spectrum: Cosmic Filaments and the Temperature of Dark Matter. *ApJ*, 762:115, January 2013. doi:10.1088/0004-637X/762/2/115.

Okumura, T., Matsubara, T., Eisenstein, D. J., et al. Large-Scale Anisotropic Correlation Function of SDSS Luminous Red Galaxies. *ApJ*, 676:889–898, April 2008. doi:10.1086/528951.

Ostriker, J. P. and Peebles, P. J. E. A Numerical Study of the Stability of Flattened Galaxies: or, can Cold Galaxies Survive? *ApJ*, 186:467–480, December 1973. doi:10.1086/152513.

Ostriker, J. P., Peebles, P. J. E., and Yahil, A. The Size and Mass of Galaxies, and the Mass of the Universe. *ApJ*, 193:L1, October 1974. doi:10.1086/181617.

Padmanabhan, N. and White, M. Constraining anisotropic baryon oscillations. *PRD*, 77: 123540, June 2008. doi:10.1103/PhysRevD.77.123540.

Park, C., Vogeley, M. S., Geller, M. J., and Huchra, J. P. Power Spectrum, Correlation Function, and Tests for Luminosity Bias in the CfA Redshift Survey. *ApJ*, 431:569, August 1994. doi:10.1086/174508.

Peacock, J. A. and Dodds, S. J. Reconstructing the Linear Power Spectrum of Cosmological Mass Fluctuations. *MNRAS*, 267:1020, April 1994. doi:10.1093/mnras/267.4.1020.

Peacock, J. A. and Heavens, A. F. Alternatives to the Press-Schechter cosmological mass function. *MNRAS*, 243:133–143, March 1990. doi:10.1093/mnras/243.1.133.

Peacock, J. A. and Smith, R. E. Halo occupation numbers and galaxy bias. *Mon. Not. Roy. Astron. Soc.*, 318:1144, 2000. doi:10.1046/j.1365-8711.2000.03779.x.

Peebles, P. J. E. *The large-scale structure of the universe*. 1980.

Peebles, P. J. E. Large-scale background temperature and mass fluctuations due to scale- invariant primeval perturbations. *ApJ*, 263:L1–L5, December 1982. doi:10.1086/183911.

Peebles, P. J. E. Tests of cosmological models constrained by inflation. *ApJ*, 284:439–444, September 1984. doi:10.1086/162425.

Peebles, P. J. E. and Groth, E. J. Statistical analysis of catalogs of extragalactic objects. V. Three- point correlation function for the galaxy distribution in the Zwicky catalog. *ApJ*, 196:1–11, February 1975. doi:10.1086/153390.

Peebles, P. J. E. and Hauser, M. G. Statistical Analysis of Catalogs of Extragalactic Objects. III. The Shane-Wirtanen and Zwicky Catalogs. *The Astrophysical Journal Supplement Series*, 28:19, November 1974. doi:10.1086/190308.

Penzias, A. A. and Wilson, R. W. A Measurement of Excess Antenna Temperature at 4080 Mc/s. *ApJ*, 142:419–421, July 1965. doi:10.1086/148307.

Percival, W. J., Baugh, C. M., Bland-Hawthorn, J., et al. The 2dF Galaxy Redshift Survey: the power spectrum and the matter content of the Universe. *MNRAS*, 327:1297–1306, November 2001. doi:10.1046/j.1365-8711.2001.04827.x.

Perlmutter, S., Aldering, G., Goldhaber, G., et al. Measurements of Ω and Λ from 42 High-Redshift Supernovae. *ApJ*, 517:565–586, June 1999. doi:10.1086/307221.

Pichon, C. and Bernardeau, F. Vorticity generation in large-scale structure caustics. *A&A*, 343:663–681, March 1999.

Planck Collaboration, Adam, R., Ade, P. A. R., et al. Planck 2015 results. I. Overview of products and scientific results. *A&A*, 594:A1, September 2016a. doi:10.1051/0004-6361/201527101.

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. Planck 2015 results. XIII. Cosmological parameters. *A&A*, 594:A13, September 2016b. doi:10.1051/0004-6361/201525830.

Pollack, J. E., Smith, R. E., and Porciani, C. Modelling large-scale halo bias using the bispectrum. *MNRAS*, 420:3469–3489, March 2012. doi:10.1111/j.1365-2966.2011.20279.x.

Pollack, J. E., Smith, R. E., and Porciani, C. A new method to measure galaxy bias. *MNRAS*, 440:555–576, May 2014. doi:10.1093/mnras/stu322.

Praton, E. A., Melott, A. L., and McKee, M. Q. The Bull's-Eye Effect: Are Galaxy Walls Observationally Enhanced? *ApJ*, 479:L15–L18, April 1997. doi:10.1086/310567.

Press, W. H. and Schechter, P. Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. *ApJ*, 187:425–438, February 1974. doi:10.1086/152650.

Pueblas, S. and Scoccimarro, R. Generation of vorticity and velocity dispersion by orbit crossing. *PRD*, 80, August 2009. doi:10.1103/PhysRevD.80.043504.

Ratra, B. and Peebles, P. J. E. Cosmological consequences of a rolling homogeneous scalar field. *PRD*, 37:3406–3427, June 1988. doi:10.1103/PhysRevD.37.3406.

Regan, D. M., Shellard, E. P. S., and Fergusson, J. R. General CMB and Primordial Trispectrum Estimation. *Phys. Rev.*, D82:023520, 2010. doi:10.1103/PhysRevD.82.023520.

Regan, D. M., Schmittfull, M. M., Shellard, E. P. S., and Fergusson, J. R. Universal Non-Gaussian Initial Conditions for N-body Simulations. *Phys. Rev.*, D86:123524, 2012. doi:10.1103/PhysRevD.86.123524.

Regan, D. M., Schmittfull, M. M., Shellard, E. P. S., and Fergusson, J. R. Universal non-Gaussian initial conditions for N-body simulations. *PRD*, 86(12):123524, December 2012. doi:10.1103/PhysRevD.86.123524.

Reid, B., Ho, S., Padmanabhan, N., et al. SDSS-III Baryon Oscillation Spectroscopic Survey Data Release 12: galaxy target selection and large-scale structure catalogues. *MNRAS*, 455:1553–1573, January 2016. doi:10.1093/mnras/stv2382.

Riess, A. G., Filippenko, A. V., Challis, P., et al. Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. AJ, 116:1009–1038, September 1998. doi:10.1086/300499.

Riess, A. G., Casertano, S., Yuan, W., et al. New Parallaxes of Galactic Cepheids from Spatially Scanning the Hubble Space Telescope: Implications for the Hubble Constant. *ApJ*, 855:136, March 2018. doi:10.3847/1538-4357/aaadb7.

Robertson, H. P. Kinematics and World-Structure. *ApJ*, 82:284, November 1935. doi:10.1086/143681.

Rubin, V. C., Ford, W. K., and Thonnard, N. Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 (R=4kpc) to UGC 2885 (R=122kpc). *ApJ*, 238:471–487, June 1980. doi:10.1086/158003.

Rubin, V. C. and Ford, J., W. Kent. Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. *ApJ*, 159:379, February 1970. doi:10.1086/150317.

Ryden, B. S. and Gramann, M. Phase Shifts in Gravitationally Evolving Density Fields. *ApJ*, 383:L33, December 1991. doi:10.1086/186234.

Saito, S., Baldauf, T., Vlah, Z., et al. Understanding higher-order nonlocal halo bias at large scales by combining the power spectrum with the bispectrum. *PRD*, 90(12): 123522, December 2014. doi:10.1103/PhysRevD.90.123522.

Sánchez, A. G., Kazin, E. A., Beutler, F., et al. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological constraints from the full shape of the clustering wedges. *MNRAS*, 433:1202–1222, August 2013. doi:10.1093/mnras/stt799.

Sánchez, A. G., Scoccimarro, R., Crocce, M., et al. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: Cosmological implications of the configuration-space clustering wedges. *MNRAS*, 464:1640–1658, January 2017. doi:10.1093/mnras/stw2443.

Sato, M. and Nishimichi, T. Impact of the non-Gaussian covariance of the weak lensing power spectrum and bispectrum on cosmological parameter estimation. *Phys. Rev.*, D87:123538, 2013. doi:10.1103/PhysRevD.87.123538.

Schmidt, F., Jeong, D., and Desjacques, V. Peak-background split, renormalization, and galaxy clustering. *PRD*, 88:023515, July 2013. doi:10.1103/PhysRevD.88.023515.

Schmittfull, M., Baldauf, T., and Seljak, U. Near optimal bispectrum estimators for large-scale structure. *PRD*, 91(4):043530, February 2015. doi:10.1103/PhysRevD.91.043530.

Schmittfull, M. M., Regan, D. M., and Shellard, E. P. S. Fast Estimation of Gravitational and Primordial Bispectra in Large Scale Structures. *Phys. Rev.*, D88(6):063512, 2013. doi:10.1103/PhysRevD.88.063512.

Scoccimarro, R. and Frieman, J. Loop Corrections in Nonlinear Cosmological Perturbation Theory. *ApJS*, 105:37, July 1996. doi:10.1086/192306.

Scoccimarro, R. Cosmological Perturbations: Entering the Nonlinear Regime. *ApJ*, 487: 1–17, September 1997. doi:10.1086/304578.

Scoccimarro, R. Transients from initial conditions: a perturbative analysis. *MNRAS*, 299: 1097–1118, October 1998. doi:10.1046/j.1365-8711.1998.01845.x.

Scoccimarro, R. The bispectrum: from theory to observations. *Astrophys. J.*, 544:597, 2000. doi:10.1086/317248.

Scoccimarro, R. A New Angle on Gravitational Clustering. In *Annals of the New York Academy of Sciences, vol. 927, issue 1, pp. 13-23*, volume 927, pages 13–23, January 2001. doi:10.1111/j.1749-6632.2001.tb05618.x.

Scoccimarro, R. Redshift-space distortions, pairwise velocities, and nonlinearities. *PRD*, 70:083007, October 2004. doi:10.1103/PhysRevD.70.083007.

Scoccimarro, R. Fast Estimators for Redshift-Space Clustering. *Phys. Rev.*, D92(8):083532, 2015. doi:10.1103/PhysRevD.92.083532.

Scoccimarro, R. and Frieman, J. Loop corrections in nonlinear cosmological perturbation theory 2. Two point statistics and selfsimilarity. *Astrophys. J.*, 473:620, 1996. doi:10.1086/178177.

Scoccimarro, R. and Sheth, R. K. PTHalos: A Fast method for generating mock galaxy distributions. *Mon. Not. Roy. Astron. Soc.*, 329:629–640, 2002. doi:10.1046/j.1365-8711.2002.04999.x.

Scoccimarro, R., Colombi, S., Fry, J. N., et al. Nonlinear Evolution of the Bispectrum of Cosmological Perturbations. *ApJ*, 496:586–604, March 1998a. doi:10.1086/305399.

Scoccimarro, R., Colombi, S., Fry, J. N., et al. Nonlinear Evolution of the Bispectrum of Cosmological Perturbations. *ApJ*, 496:586–604, March 1998b. doi:10.1086/305399.

Scoccimarro, R., Couchman, H. M. P., and Frieman, J. A. The Bispectrum as a Signature of Gravitational Instability in Redshift Space. *ApJ*, 517:531–540, June 1999a. doi:10.1086/307220.

Scoccimarro, R., Zaldarriaga, M., and Hui, L. Power Spectrum Correlations Induced by Nonlinear Clustering. *ApJ*, 527:1–15, December 1999b. doi:10.1086/308059.

Scoccimarro, R., Feldman, H. A., Fry, J. N., and Frieman, J. A. The Bispectrum of IRAS Redshift Catalogs. *ApJ*, 546:652–664, January 2001. doi:10.1086/318284.

Scoccimarro, R., Sheth, R. K., Hui, L., and Jain, B. How many galaxies fit in a halo? Constraints on galaxy formation efficiency from spatial clustering. *Astrophys. J.*, 546: 20–34, 2001. doi:10.1086/318261.

Scoccimarro, R., Sefusatti, E., and Zaldarriaga, M. Probing primordial non-Gaussianity with large - scale structure. *Phys. Rev.*, D69:103513, 2004. doi:10.1103/PhysRevD.69.103513.

Sefusatti, E. *Probing fundamental physics with large-scale structure: From galaxy formation to inflation*. PhD thesis, New York University, New York, USA, December 2005.

Sefusatti, E. and Komatsu, E. The bispectrum of galaxies from high-redshift galaxy surveys: Primordial non-Gaussianity and non-linear galaxy bias. *Phys. Rev.*, D76: 083004, 2007. doi:10.1103/PhysRevD.76.083004.

Sefusatti, E. and Scoccimarro, R. Galaxy bias and halo-occupation numbers from large-scale clustering. *PRD*, 71:063001, March 2005. doi:10.1103/PhysRevD.71.063001.

Sefusatti, E., Crocce, M., Pueblas, S., and Scoccimarro, R. Cosmology and the bispectrum. *PRD*, 74:023522, July 2006. doi:10.1103/PhysRevD.74.023522.

Sefusatti, E., Crocce, M., and Desjacques, V. The Halo Bispectrum in N-body Simulations with non-Gaussian Initial Conditions. *Mon. Not. Roy. Astron. Soc.*, 425:2903, 2012. doi:10.1111/j.1365-2966.2012.21271.x.

Seljak, U. and Zaldarriaga, M. A Line-of-Sight Integration Approach to Cosmic Microwave Background Anisotropies. *ApJ*, 469:437, October 1996. doi:10.1086/177793.

Seljak, U. Analytic model for galaxy and dark matter clustering. *Mon. Not. Roy. Astron. Soc.*, 318:203, 2000. doi:10.1046/j.1365-8711.2000.03715.x.

Sellentin, E. and Heavens, A. F. Parameter inference with estimated covariance matrices. *Mon. Not. Roy. Astron. Soc.*, 456(1):L132–L136, 2016. doi:10.1093/mnrasl/slv190.

Seo, H.-J., Sato, M., Dodelson, S., Jain, B., and Takada, M. Re-capturing cosmic information. *Astrophys. J.*, 729:L11, 2011. doi:10.1088/2041-8205/729/1/L11.

Shane, C. D. and Wirtanen, C. A. *Pub. Lick Obs.*, 22:part 1, 1967.

Sheth, R. K. and Tormen, G. Large-scale bias and the peak background split. *MNRAS*, 308:119–126, September 1999. doi:10.1046/j.1365-8711.1999.02692.x.

Sheth, R. K. and Tormen, G. An excursion set model of hierarchical clustering: ellipsoidal collapse and the moving barrier. *MNRAS*, 329:61–75, January 2002. doi:10.1046/j.1365-8711.2002.04950.x.

Sheth, R. K., Mo, H. J., and Tormen, G. Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes. *MNRAS*, 323:1–12, May 2001. doi:10.1046/j.1365-8711.2001.04006.x.

Silk, J. Challenges in Cosmology from the Big Bang to Dark Energy, Dark Matter and Galaxy Formation. In *14th International Symposium on Nuclei in the Cosmos (NIC2016)*, page 010101, January 2017. doi:10.7566/JPSCP.14.010101.

Simonović, M., Baldauf, T., Zaldarriaga, M., Carrasco, J. J., and Kollmeier, J. A. Cosmological perturbation theory using the FFTLog: formalism and connection to QFT loop integrals. *Journal of Cosmology and Astro-Particle Physics*, 2018:030, April 2018. doi:10.1088/1475-7516/2018/04/030.

Slepian, Z., Eisenstein, D. J., Beutler, F., et al. The large-scale three-point correlation function of the SDSS BOSS DR12 CMASS galaxies. *MNRAS*, 468:1070–1083, June 2017a. doi:10.1093/mnras/stw3234.

Slepian, Z., Eisenstein, D. J., Brownstein, J. R., et al. Detection of baryon acoustic oscillation features in the large-scale three-point correlation function of SDSS BOSS DR12 CMASS galaxies. *MNRAS*, 469:1738–1751, August 2017b. doi:10.1093/mnras/stx488.

Smith, R. E. Covariance of cross-correlations: towards efficient measures for large-scale structure. *MNRAS*, 400:851–865, December 2009. doi:10.1111/j.1365-2966.2009.15490.x.

Smith, R. E. and Watts, P. I. R. Triaxial haloes, intrinsic alignments and the dark matter power spectrum. *MNRAS*, 360:203–215, June 2005. doi:10.1111/j.1365-2966.2005.09053.x.

Smith, R. E., Peacock, J. A., Jenkins, A., et al. Stable clustering, the halo model and nonlinear cosmological power spectra. *Mon. Not. Roy. Astron. Soc.*, 341:1311, 2003. doi:10.1046/j.1365-8711.2003.06503.x.

Smith, R. E., Scoccimarro, R., and Sheth, R. K. Scale dependence of halo and galaxy bias: Effects in real space. *PRD*, 75(6):063512–+, March 2007. doi:10.1103/PhysRevD.75.063512.

Smith, R. E., Reed, D. S., Potter, D., et al. Precision cosmology in muddy waters: cosmological constraints and N-body codes. *MNRAS*, 440:249–268, May 2014. doi:10.1093/mnras/stu272.

Smith, R. E. and Marian, L. Towards optimal estimation of the galaxy power spectrum. *MNRAS*, 454:1266–1289, December 2015. doi:10.1093/mnras/stv2042.

Smith, R. E. and Marian, L. What is the optimal way to measure the galaxy power spectrum? *MNRAS*, 457:4285–4290, April 2016. doi:10.1093/mnras/stw282.

Smith, R. E., Scoccimarro, R., and Sheth, R. K. Motion of the acoustic peak in the correlation function. *PRD*, 77:043525, February 2008. doi:10.1103/PhysRevD.77.043525.

Smoot, G. F., Bennett, C. L., Kogut, A., et al. Structure in the COBE Differential Microwave Radiometer First-Year Maps. *ApJ*, 396:L1, September 1992. doi:10.1086/186504.

Soda, J. and Suto, Y. Nonlinear Gravitational Evolution of Phases and Amplitudes in One- dimensional Cosmological Density Fields. *ApJ*, 396:379, September 1992. doi:10.1086/171726.

Song, Y.-S., Taruya, A., and Oka, A. Cosmology with anisotropic galaxy clustering from the combination of power spectrum and bispectrum. *Journal of Cosmology and Astro-Particle Physics*, 2015:007, August 2015. doi:10.1088/1475-7516/2015/08/007.

Spergel, D., Gehrels, N., Baltay, C., et al. Wide-Field InfrarRed Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA 2015 Report. *ArXiv e-prints*, art. arXiv:1503.03757, March 2015.

Spergel, D. N. The dark side of cosmology: Dark matter and dark energy. *Science*, 347: 1100–1102, March 2015. doi:10.1126/science.aaa0980.

Springel, V. The cosmological simulation code GADGET-2. *MNRAS*, 364:1105–1134, December 2005. doi:10.1111/j.1365-2966.2005.09655.x.

Starobinsky, A. A. Dynamics of phase transition in the new inflationary universe scenario and generation of perturbations. *Physics Letters B*, 117:175–178, November 1982. doi:10.1016/0370-2693(82)90541-X.

Suto, Y. and Sasaki, M. Quasinonlinear theory of cosmological self-gravitating systems. *Physical Review Letters*, 66:264–267, January 1991. doi:10.1103/PhysRevLett.66.264.

Szalay, A. S. Constraints on the biasing of density fluctuations. *ApJ*, 333:21–23, October 1988. doi:10.1086/166721.

Szapudi, I. and Colombi, S. Cosmic Error and Statistics of Large-Scale Structure. *ApJ*, 470:131, October 1996. doi:10.1086/177855.

Takada, M. and Bridle, S. Probing dark energy with cluster counts and cosmic shear power spectra: including the full covariance. *New J. Phys.*, 9:446, 2007. doi:10.1088/1367-2630/9/12/446.

Takada, M., Ellis, R. S., Chiba, M., et al. Extragalactic science, cosmology, and Galactic archaeology with the Subaru Prime Focus Spectrograph. *Publications of the Astronomical Society of Japan*, 66:R1, February 2014. doi:10.1093/pasj/pst019.

Takahashi, R., Yoshida, N., Takada, M., et al. Simulations of Baryon Acoustic Oscillations. II. Covariance Matrix of the Matter Power Spectrum. *ApJ*, 700:479–490, July 2009. doi:10.1088/0004-637X/700/1/479.

Takahashi, R., Yoshida, N., Takada, M., et al. Non-Gaussian Error Contribution to Likelihood Analysis of the Matter Power Spectrum. *Astrophys. J.*, 726:7, 2011. doi:10.1088/0004-637X/726/1/7.

Taruya, A., Nishimichi, T., and Saito, S. Baryon acoustic oscillations in 2D: Modeling redshift-space power spectrum from perturbation theory. *PRD*, 82:063522, September 2010. doi:10.1103/PhysRevD.82.063522.

Taruya, A., Bernardeau, F., Nishimichi, T., and Codis, S. Direct and fast calculation of regularized cosmological power spectrum at two-loop order. *PRD*, 86:103528, November 2012. doi:10.1103/PhysRevD.86.103528.

Taruya, A., Nishimichi, T., and Bernardeau, F. Precision modeling of redshift-space distortions from a multipoint propagator expansion. *PRD*, 87:083509, April 2013. doi:10.1103/PhysRevD.87.083509.

Tegmark, M., Taylor, A. N., and Heavens, A. F. Karhunen-Loève Eigenvalue Problems in Cosmology: How Should We Tackle Large Data Sets? *ApJ*, 480:22–35, May 1997.

Tegmark, M., Eisenstein, D. J., Strauss, M. A., et al. Cosmological constraints from the SDSS luminous red galaxies. *PRD*, 74(12):123507, December 2006.

Tegmark, M. Measuring Cosmological Parameters with Galaxy Surveys. *PRL*, 79:3806–3809, November 1997. doi:10.1103/PhysRevLett.79.3806.

Tellarini, M., Ross, A. J., Tasinato, G., and Wands, D. Galaxy bispectrum, primordial non-Gaussianity and redshift space distortions. *JCAP*, 1606(06):014, 2016. doi:10.1088/1475-7516/2016/06/014.

Tinker, J. L., Robertson, B. E., Kravtsov, A. V., et al. The Large-scale Bias of Dark Matter Halos: Numerical Calibration and Model Tests. *ApJ*, 724:878–886, December 2010. doi:10.1088/0004-637X/724/2/878.

Turner, M. S., Steigman, G., and Krauss, L. M. Flatness of the Universe: Reconciling Theoretical Prejudices with Observational Data. *PRL*, 52:2090–2093, June 1984. doi:10.1103/PhysRevLett.52.2090.

van Uitert, E., Joachimi, B., Joudaki, S., et al. KiDS+GAMA: cosmology constraints from a joint analysis of cosmic shear, galaxy-galaxy lensing, and angular clustering. *MNRAS*, 476:4662–4689, June 2018. doi:10.1093/mnras/sty551.

Verde, L., Heavens, A. F., Percival, W. J., et al. The 2dF Galaxy Redshift Survey: the bias of galaxies and the density of the Universe. *MNRAS*, 335:432–440, September 2002. doi:10.1046/j.1365-8711.2002.05620.x.

Vishniac, E. T. The dynamic and gravitational instabilities of spherical shocks. *ApJ*, 274:152–167, November 1983. doi:10.1086/161433.

Wagner, C., Schmidt, F., Chiang, C. T., and Komatsu, E. Separate universe simulations. *MNRAS*, 448:L11–L15, March 2015. doi:10.1093/mnrasl/slu187.

Walker, A. G. On Milne's Theory of World-Structure. *Proceedings of the London Mathematical Society*, 42:90–127, January 1937. doi:10.1112/plms/s2-42.1.90.

Walker, M. G. and Peñarrubia, J. A Method for Measuring (Slopes of) the Mass Profiles of Dwarf Spheroidal Galaxies. *ApJ*, 742:20, November 2011. doi:10.1088/0004-637X/742/1/20.

Watts, P., Coles, P., and Melott, A. The Universal Behavior of Phase Correlations in Nonlinear Gravitational Clustering. *ApJL*, 589:L61–L64, June 2003. doi:10.1086/376351.

Weinberg, D. H., Mortonson, M. J., Eisenstein, D. J., et al. Observational probes of cosmic acceleration. *Phys. Rep.* , 530:87–255, September 2013. doi:10.1016/j.physrep.2013.05.001.

Weinberg, S. The cosmological constant problem. *Reviews of Modern Physics*, 61:1–23, January 1989. doi:10.1103/RevModPhys.61.1.

Weinberg, S. *Cosmology*. 2008.

Welling, Y., van der Woude, D., and Pajer, E. Lifting Primordial Non-Gaussianity Above the Noise. *JCAP*, 1608(08):044, 2016. doi:10.1088/1475-7516/2016/08/044.

Wolstenhulme, R., Bonvin, C., and Obreschkow, D. Three-point Phase Correlations: A New Measure of Non-linear Large-scale Structure. *ApJ*, 804:132, May 2015. doi:10.1088/0004-637X/804/2/132.

York, D. G., Adelman, J., Anderson, J., John E., et al. The Sloan Digital Sky Survey: Technical Summary. AJ, 120:1579–1587, September 2000. doi:10.1086/301513.

Yu, J. T. and Peebles, P. J. E. Superclusters of Galaxies? *ApJ*, 158:103, October 1969. doi:10.1086/150175.

Zehavi, I., Zheng, Z., Weinberg, D. H., et al. Galaxy Clustering in the Completed SDSS Redshift Survey: The Dependence on Color and Luminosity. *ApJ*, 736:59, July 2011. doi:10.1088/0004-637X/736/1/59.

Zel'Dovich, Y. B. Cosmological Constant and Elementary Particles. *Soviet Journal of Experimental and Theoretical Physics Letters*, 6:316, November 1967.

Zel'dovich, Y. B. Special Issue: the Cosmological Constant and the Theory of Elementary Particles. *Soviet Physics Uspekhi*, 11:381–393, March 1968. doi:10.1070/PU1968v011n03ABEH003927.

Zel'dovich, Y. B. Gravitational instability: An approximate theory for large density perturbations. *A&A*, 500:13–18, June 2009.

Zentner, A. R. The Excursion Set Theory of Halo Mass Functions, Halo Clustering, and Halo Growth. *International Journal of Modern Physics D*, 16:763–815, January 2007. doi:10.1142/S0218271807010511.

Zheng, Y., Zhang, P., and Jing, Y. Determination of the large scale volume weighted halo velocity bias in simulations. *PRD*, 91:123512, June 2015. doi:10.1103/PhysRevD.91.123512.

Zheng, Z., Coil, A. L., and Zehavi, I. Galaxy Evolution from Halo Occupation Distribution Modeling of DEEP2 and SDSS Galaxy Clustering. *ApJ*, 667:760–779, October 2007. doi:10.1086/521074.

Zu, Y. and Mandelbaum, R. Mapping stellar content to dark matter haloes using galaxy clustering and galaxy-galaxy lensing in the SDSS DR7. *MNRAS*, 454:1161–1191, December 2015. doi:10.1093/mnras/stv2062.

Zwicky, F. Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta*, 6: 110–127, January 1933.

Zwicky, I. F. On the Clustering of Nebulae. *ApJ*, 95:555, May 1942. doi:10.1086/144422.